ISCR

Institute for Scientific Computing Research

Fiscal Year 2001 Annual Report

http://www.llnl.gov/casc/iscr/

Lawrence Livermore National Laboratory
P.O. Box 808, L-561, Livermore, CA 94551

CASC

Center for Applied Scientific Computing

UNIVERSITY RELATIONS PROGRAM

The University Relations Program (URP) encourages collaborative research between Lawrence Livermore National Laboratory (LLNL) and the University of California campuses. The Institute for Scientific Computing Research (ISCR) actively participates in such collaborative research, and this report details the Fiscal Year 2001 projects jointly served by URP and ISCR. For a full discussion of all URP projects in FY 2001, please request a copy of the URP FY 2001 Annual Report by contacting

Lawrence Livermore National Laboratory
Edie Rock, University Relations Program
P. O. Box 808, L-413
Livermore, CA 94551

UCRL-LR-133866-01

Work performed under the auspices of the U.S. Department of Energy by University of California Lawrence Livermore National Laboratory under Contract W-7405-ENG-48.

Contents

The Mission of the ISCR

The Institute for Scientific Computing Research (ISCR) at Lawrence Livermore National Laboratory is jointly administered by the Center for Applied Scientific Computing (CASC) and the University Relations Program (URP), and this joint relationship expresses its mission. An extensively externally networked ISCR cost-effectively expands the level and scope of national computational science expertise available to the laboratory through CASC. The URP, with its infrastructure for managing five institutes and numerous educational programs at LLNL, assumes much of the logistical burden that is unavoidable in bridging the laboratory's internal computational research environment with that of the academic community.

As large-scale simulations on the parallel platforms of DOE's Accelerated Strategic Computing Initiative become increasingly important to the overall mission of LLNL, the role of the ISCR expands in importance, accordingly.

Relying primarily on non-permanent staffing, the ISCR complements laboratory research in areas of the computer and information sciences that are needed at the frontier of laboratory missions. The ISCR strives to be the "eyes and ears" of the laboratory in the computer and information sciences, in keeping the laboratory aware of and connected to important external advances. It also attempts to be "feet and hands," in carrying those advances into the laboratory and incorporating them into practice. In addition to conducting research, the ISCR provides continuing education opportunities to laboratory personnel, in the form of on-site workshops taught by experts on novel software or hardware technologies.

The ISCR also seeks to influence the research community external to the laboratory to pursue laboratory-related interests and to train the workforce that will be required by the laboratory. Part of the performance of this function is interpreting to the external community appropriate (unclassified) aspects of the laboratory's own contributions to the computer and information sciences—contributions that its unique mission and unique resources give it a unique opportunity and responsibility to make.

Of the three principal means of packaging scientific ideas for transfer—people, papers, and software—experience suggests that the most effective means is people. The programs of the ISCR are therefore people-intensive.

Finally, the ISCR, together with CASC, confers an organizational identity on the burgeoning computer and information sciences research activity at LLNL and serves as a point of contact within the laboratory for computer and information scientists from outside.

Institute for Scientific Computing Research
Fiscal Year 2001 Director's Report

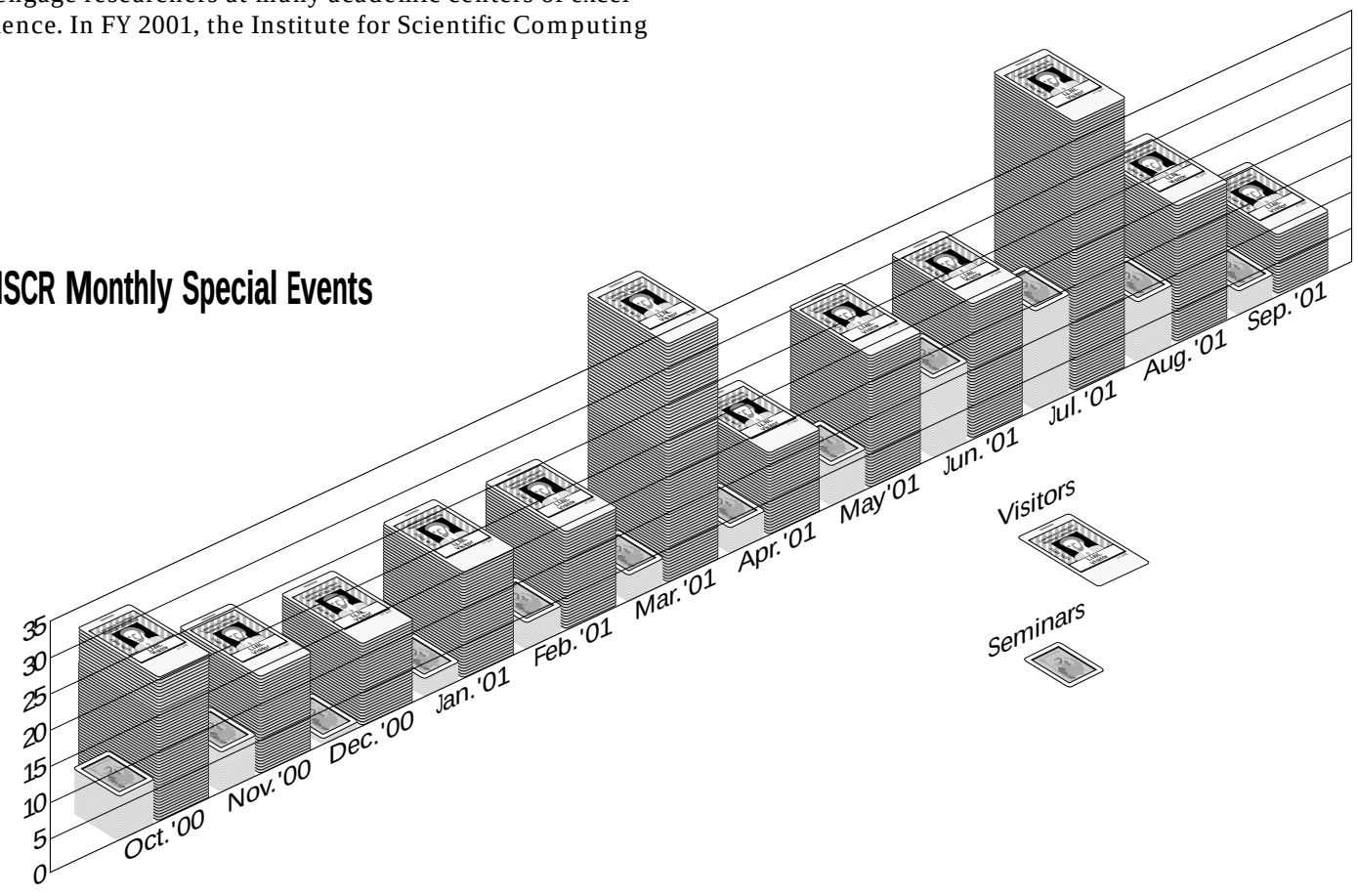
Large-scale scientific computation, and all of the disciplines that support it and help to validate it, have been placed at the focus of Lawrence Livermore National Laboratory by the Advanced Simulation and Computing (ASCI) program and more recently by DOE's Scientific Discovery through Advanced Computing (SciDAC) initiative. The Laboratory operates the computer with the highest peak performance in the world and has undertaken some of the largest and most compute-intensive simulations ever performed. Scientific simulation was the featured discipline at the Laboratory's 2001 public "Science Day" program. However, computers at architectural extremes are notoriously difficult to use efficiently. Furthermore, each successful terascale simulation only points out the need for much better ways of interacting with the resulting data.

Advances in scientific computing research have therefore never been more vital to the core missions of the Laboratory than at present. Computational science is evolving so rapidly along every one of its research fronts that to remain on the leading edge the Laboratory must engage researchers at many academic centers of excellence. In FY 2001, the Institute for Scientific Computing Research (ISCR) has served as one of the Laboratory's main bridges to the academic community in the form of collaborative subcontracts, visiting faculty, student internships, workshops, and an active seminar series.

ISCR research participants are integrated into the Laboratory's Center for Applied Scientific Computing (CASC), which, in turn, addresses computational challenges arising throughout the Laboratory. Administratively, the ISCR flourishes under the Laboratory's University Relations Program (URP). Together with the other four Institutes of the URP, it navigates a course that allows the Laboratory to benefit from academic exchanges while preserving national security. While it is difficult to operate an academic-like research enterprise within the context of a national security laboratory, the results declare the challenges well met and worth the continued effort.

Fiscal year 2001 was the second full year under Acting Director David Keyes. Keyes, the Richard F. Barry

ISCR Monthly Special Events



Professor of Mathematics & Statistics at Old Dominion University and an ISCR faculty participant since October 1997, dedicated one-third of his time to the technical program of the ISCR. Jill Dunaway continued as the full-time Institute Administrator. Terry Garrigan, Emma Horcabas, and Leslie Bills all assisted with the large visitor and summer program. With a reorganization of the Computation Directorate at LLNL in November 2001, Dunaway moved on to the administration of CASC, itself, and the ISCR is most delighted that Linda Bodtker has come on board as the new Administrator for FY02.

In FY 200l, we continued our Institute for Terascale Simulation Lecture Series, featuring visits from Ingrid Daubechies, David Johnson, Michael Norman, Charles Peskin, Eugene Spafford, and Andries Van Dam. A special section of this annual report is devoted to the abstracts and biosketches of these distinguished lecturers. The ITS Lectures typically draw two to three hundred people from around the Laboratory and surrounding scientific community. They are archived on video and available at the LLNL Technical Library. We plan to continue this series with approximately six "movers and shakers" in high-end simulation and its enabling technologies per year.

In February, the ISCR hosted the Bay Area Scientific Computing Day, an annual gathering designed to strengthen ties between scientific computing researchers throughout the Bay Area, featuring talks by students, post-docs, and senior researchers, and drawing 120 participants.

In a series of twelve lectures throughout the spring of 2001, sabbatical visitor Professor Omar Ghattas of Carnegie Mellon University, a leader in the field of optimization subject to large-scale constraints of partial differential equation type, gave a short course to approximately twenty regular attendees on optimization techniques in computational science.

In early April, the ISCR co-sponsored three international conferences held off-site with significant technical leadership from permanent CASC staff and ISCR affiliates, beginning with the annual Copper Mountain Conference, in Copper Mountain, Colorado. The 2001 meeting was devoted to Multigrid Methods. Four members of the CASC scientific staff presented papers, as did nineteen of the academic collaborators of the ISCR. Van Emden Henson of CASC presented a multigrid tutorial on the opening day.

The ISCR also continued its role in promoting scientific aspects of data mining, with co-sponsorship of the Third Workshop on Mining Scientific Data Sets. This one-day workshop was held in Chicago, in conjunction with a data mining workshop of larger scope, organized by the Society for Industrial and Applied Mathematics (SIAM).

In late April, the ISCR co-sponsored another international conference dedicated to Preconditioning Techniques for Large Sparse Matrix Problems in Industrial Applications in Tahoe City, California.

In June, with the advent of our large student summer program and sponsorship from the Defense Programs office of DOE HQ, we ramped up our second annual Internships in Terascale Simulation Technology tutorial series. The tutors included CASC's textbook authors, John May and Van Emden Henson, CASC computational mathematicians David Brown and Carol Woodward, CASC computer scientists Terence Critchlow and Gary Kumfert, three visiting faculty, and the ISCR Director. Though intended for students, permanent CASC researchers attended an occasional subseries of the lectures.

Also in June, under the direction of CASC scientists Jim Jones and Rob Falgout, the ISCR organized a three-day Workshop on Solution Methods for Large-scale Nonlinear Problems in Livermore.

In July, the ISCR organized in Livermore a three-day Workshop on Object-Oriented and Component Technology for Scientific Computing, under the direction of CASC scientists Scott Kohn and Gary Kumfert.

Throughout FY 2001, the ISCR brought to the laboratory a vigorous contingent of post-docs, faculty visitors, and students. There were 27 faculty visitors in residence for more than just a seminar visit – for a week to a semester. Eight post-docs made the ISCR their home this past year. We also had 44 students in residence, mostly for 8–10 weeks of the summer, but several of them for a semester or a full year. Each of these students was in a research relationship with one of CASC's approximately ninety full-time technical staff.

The pages of this report summarize the activities of the faculty members, post-doctoral researchers, students, and guests from industry and other laboratories who participated in LLNL's computational mission under the auspices of the ISCR during FY 2001. Altogether, the ISCR hosted 223 visits from 182 different visitors, who gave a

total of 73 seminars on site. The vast majority of the visitors were from academia, with 15% from industry and 15% from other laboratories. Visitors from outside of the United States made up 10% of the total. The histogram on page 4 charts the numbers of visitors and seminars as a function of the month of the fiscal year.

Most of the material of this annual report comes directly from the visitors and principal investigators of the projects being reported, who selected formats convenient for their purposes. We thank Whitney Lacy for her editorial work and Dan Moore of the Technical Information Division of LLNL for his graphic artistry in producing an easily navigated and visually pleasing document.

We hope that you enjoy examining this report on the ISCR's diverse activities in FY 2001. For further information about the Institute, please contact us at the address below. Inquiries about how you might enhance the on-going FY 2002 program at the ISCR, or beyond, are welcome.

David Keyes

David E. Keyes
Acting Director
http://www.llnl.gov/casc/people/keyes
dekeyes@llnl.gov
925-422-1325

Jill Dunaway
Administrator
http://www.llnl.gov/casc/people/dunaway
dunaway4@llnl.gov
925-422-7132

ISCR
Institute for Scientific Computing Research

Lawrence Livermore National Laboratory
P. O. Box 808, L-561
Livermore, CA 94551

http://www.llnl.gov/casc/iscr/

ISCR Fiscal Year 2001 in Review

FY 2001 Seminar Series (in reverse chronological order)

Mark Mitchell, CodeSourcery, LLC (Seminar) ..... September 28, 2001
Don Schwendeman, Rensselaer Polytechnic Institute ..... September 24-28, 2001
Sutanu Sarkar, University of California, Davis (Seminar) ..... September 26-27, 2001
Ken Joy, University of California, Davis ..... September 24-26, 2001
James Kuo, Livermore Software Technology Corporation ..... September 21, 2001
Gerhard Starke, University of Hannover ..... September 10 & 14, 2001
Doug Swesty, State University of New York, Stony Brook ..... September 5 & 7, 2001
Ken Joy, University of California, Davis ..... September 4-7, 2001
Preethy Vaidyanathan, University of California, Santa Cruz ..... September 6, 2001
Gerhard Wellein, University of Erlangen-Nuernberg (Seminar) ..... September 6, 2001
Irad Yavneh, Technion-Israel Institute of Technology (Seminar) ..... September 1, 2001
Steven Reinhardt, University of Michigan (Seminar) ..... August 31, 2001
Johannes Kraus, University of Leoben (Seminar) ..... August 31, 2001
Tim Chartier, University of Washington ..... August 30, 2001
Michael Gertz, University of California, Davis ..... August 30, 2001
Ken Joy, University of California, Davis ..... August 27-30, 2001
Peter Arbenz, Swiss Federal Institue of Technology (Seminar) ..... August 28, 2001
Tamara Munzner, Compaq Systems Research Center (Seminar) ..... August 24, 2001
Tom Manteuffel, University of Colorado ..... August 22-24, 2001
Steve McCormick, University of Colorado ..... August 22-24, 2001
Brent Lindquist, State University of New York (Seminar) ..... August 21-23, 2001
Philip Roe, University of Michigan (Seminar) ..... August 21-22, 2001
Todd Kermit, Zephyr-TEC ..... August 21, 2001
Linda Petzold, University of California, Santa Barbara ..... August 13-17, 2001
Sergey Kirshner, University of California, Irvine ..... August 16, 2001
Kim Yates, ET International ..... August 16, 2001
Dan Cacuci, University of California, Berkeley ..... August 15, 2001
Paul Hovland, Argonne National Laboratory ..... August 15, 2001
Levanto Schachter, University of California, Davis (Seminar) ..... August 13, 2001
Sharad Mehrotra, University of California, Irvine (Seminar) ..... August 10, 2001
Michael Mascagni, Florida State University (Seminar) ..... August 8-10, 2001
Andrew Knyazev, University of Colorado, Denver ..... August 2-3, 2001
Jesus Labarta, European Center for Parallelism of Barcelona (Seminar) ..... July 27, 2001
Frederick Wong, University of California, Berkeley ..... July 27, 2001
David Caliga, SRC Computers, Inc. (Seminar) ..... July 25 & 27, 2001
Alan Sussman, University of Maryland, (Seminar) ..... July 25-26, 2001
Jens Schmidt, University of Colorado, Boulder ..... July 23-27, 2001
Thomas Russell, University of Colorado, Denver (Seminar) ..... July 23-24, 2001
Mark Bolas, Fakespace Labs, Inc. ..... July 20, 2001
Daniel Carr, Fakespace Labs, Inc. ..... July 20, 2001
Ian McDowell, Fakespace Labs, Inc. ..... July 20, 2001

Patrick Worley, Oak Ridge National Laboratory (Seminar) .....July 19-20, 2001
Jeff Hollingsworth, University of Maryland, College Park July19, 2001
Randal Burns, IBM Almaden Research Center (Seminar) .....July 18, 2001
Jessica Masters, University of California, Santa Cruz .....July 17, 2001
Cal Ribbens, Virginia Polytechnic Institute .....July 16-26, 2001
Robert Ansell-Bell, University of Oregon .....July 16-17, 2001
Ignacio Llorente, Complutense University (Seminar) .....July 16-17, 2001
Allen Malony, University of Oregon (Seminar) .....July 16-17, 2001
Byung Lee, University of Vermont .....July 15-18, 2001
Alex Pothen, Old Dominion University .....July 13-23, 2001
Arthur Kordon, Dow Chemical Company (Seminar) .....July 13, 2001
Michelle Hribar, Pacific University (Seminar) .....July 12, 2001
Rossen Dimitrov, MPI Software Technologies Inc. .....July 9-10, 2001
David Leimbach, MPI Software Technologies Inc. .....July 9-10, 2001
Anthony Skjellum, MPI Software Technologies Inc. .....July 9-10, 2001
James O'Brien, University of California, Berkeley (Seminar) .....July 9, 2001
Daniel Boley, University of Minnesota (Seminar) .....July 6, 2001
Omar Ghattas, Carnegie Mellon University (Seminar) .....July 2, 2001
Anne Greenbaum, University of Washington (Seminar) .....June 25-July 20, 2001
Tim Kelley, North Carolina State University .....June 25-26, 2001
Randall LeVeque, University of Washington (Seminar), ...June 25-26, 2001
Bernd Hamann, University of California, Davis .....June 22, 2001
Pat Hanrahan, Stanford University .....June 22, 2001
Michael Holst, University of California, San Diego .....June 22, 2001
Klaus Stueben, GMD-Forschungszentrum Informationstechnik GmbH .....June 5-20, 2001
Eric Lum, University of California, Davis .....June 19, 2001
Eric Shaffer, University of Illinois, Urbana-Champaign (Seminar) ..... June 19, 2001
Kyle Gallivan, Florida State University .....June 2-18, 2001
Achi Brandt, Weizmann Institute of Science (Seminar) .....June 14-15, 2001
Kees Oosterlee, German National Research Laboratory (Seminar) .....June 11-15, 2001
Luiz De Rose, IBM TJ Watson Research Center (Seminar) .....June 12-14, 2001
Phillip Gibbons, Bell Labs (Seminar) .....June 8, 2001
Heinz Otto Kreiss, University of California, Los Angeles .....May 31-June 7, 2001
Jeff Gibson, Stanford University (Seminar) .....June 6, 2001
Martin Schulz, Technical University of Munich (Seminar) .....June 4, 2001
Nancy Tran, University of Illinois, Urbana-Champaign (Seminar) .....June 4, 2001
Jarek Rossignac, Georgia Institute of Technology (Seminar) .....May 31, 2001
Jack Snoeyink, University of North Carolina, Chapel Hill (Seminar) .....May 31, 2001
Richard Strelitz, Los Alamos National Laboratory .....May 31, 2001
Kwai Lam Wong, University of Tennessee, Knoxville .....May 30-31, 2001
Michael Minion, University of North Carolina (Seminar), ...May 22, 2001
Gregory Miller, University of California, Davis .....May 18, 2001
Rolf Rabenseifner, University of Stuttgart (Seminar).....May 17-18, 2001
Marian Brezina, University of Colorado .....May 14-18, 2001
Timothy Campbell, University of Arizona .....May 16, 2001

Philip Colella, Lawrence Berkeley National Laboratory (Seminar) ..... May 16, 2001
Homer Walker, Worcester Polytechnic Institute ..... May 14-16, 2001
Richard Barrett, Los Alamos National Laboratory (Seminar) ..... May 15, 2001
Rajeev Rastogi, Bell Labs ..... May 11, 2001
Stefan Lang, University of Heidelberg (Seminar) ..... May 9, 2001
Sandra Nagele, University of Heidelberg (Seminar) ..... May 9, 2001
Achim Gordner, University of Heidelberg (Seminar) ..... May 8, 2001
Gabriel Wittum, University of Heidelberg (Seminar) ..... May 8, 2001
Said Elghobashi, University of California, Irvine (Seminar) ..... May 4, 2001
Martin Bertram, University of Utah ..... April 30-May 3, 2001
Charles Hansen, University of Utah ..... May 3, 2001
Owe Axelsson, University of Nijmegen (Seminar) ..... April 15-May 2, 2001
David Lowenthal, University of Georgia (Seminar) ..... April 27, 2001
Doug Swesty, State University of New York, Stony Brook (Seminar) ..... April 26-27, 2001
Alex Pothen, Old Dominion University ..... April 24, 2001
Tanya Vassilevska, Texas A&M University ..... April 17-23, 2001
David Butler, Limit Point Systems ..... April 13, 2001
Martin Bertram, University of Utah ..... April 10-15, 2001
Boris Diskin, ICASE, NASA Langley (Seminar) ..... April 9-11, 2001
Oleg Diyankov, STRELA Open Computer Center (Seminar) ..... April 9, 2001
Yuriko Renardy, Virginia Polytechnic Institute (Seminar) ..... April 9, 2001
Sanith Wijesinghe, Massachusetts Institute of Technology ..... March 19-April 6, 2001
Kyle Gallivan, Florida State University ..... March 29-30, 2001
Yousuff Hussaini, Florida State University ..... March 29-30, 2001
Martin Bertram, University of Utah ..... March 20-25, 2001
Aleksander Slominski, Indiana University ..... March 21-24, 2001
Randall Bramley, Indiana University ..... March 22, 2001
Byung Lee, University of Vermont ..... March 18-22, 2001
Bertil Gustafsson, Stanford University (Seminar) ..... March 20, 2001
Steve McCormick, University of Colorado ..... March 14-20, 2001
Timo Bremer, University of California, Davis ..... March 19, 2001
Bernd Hamann, University of California, Davis ..... March 19, 2001
Tanya Vassilevska, Texas A&M University ..... March 8-18, 2001
Ken Joy, University of California, Davis ..... March 16, 2001
Oscar Bruno, California Institute of Technology (Seminar) ..... March 14, 2001
McKay Hyde, California Institute of Technology ..... March 14, 2001
Michael McCracken, Penn State University ..... March 6, 2001
Andrew Knyazev, University of Colorado, Denver (Seminar) ..... February 25-March 3, 2001
Bryan Biegel, NASA Ames Research Center ..... March 2, 2001
William Van Dalsem, NASA Ames Research Center ..... March 2, 2001
David Ellsworth, NASA Ames Research Center ..... March 2, 2001
William Feiereisen, NASA Ames Research Center ..... March 2, 2001
Bryan Green, NASA Ames Research Center ..... March 2, 2001
Christopher Henze, NASA Ames Research Center ..... March 2, 2001
Darold Massaro, NASA Ames Research Center ..... March 2, 2001

Patrick Moran, NASA Ames Research Center .....March 2, 2001
Scott Richardson, NASA Ames Research Center .....March 2, 2001
Guy Russell, NASA Ames Research Center .....March 2, 2001
Velvin Watson, NASA Ames Research Center .....March 2, 2001
Eugene Tu, NASA Ames Research Center .....March 2, 2001
Lori Freitag, Argonne National Laboratory .....February 28, 2001
Eric de Sturler, IBM TJ Watson Research Center .....February 26-28, 2001
Padhraic Smyth, University of California, Irvine .....February 22, 2001
Esmond Ng, Lawrence Berkeley National Laboratory. .....February 21, 2001
Rajesh Rawat, University of Utah .....February 20-21, 2001
Paul Saylor, University of Illinois, Urbana-Champaign .....February 16, 2001
John Harer, Duke University (Seminar) .....February 16, 2001
Doug Swesty, State University of New York at Stony Brook .....February 15-16, 2001
Martin Bertram, University of Utah .....February 13-23, 2001
Donald Estep, Colorado State University (Seminar) .....February 9, 2001
Tom Manteuffel, University of Colorado .....February 5-9, 2001
Justin Koo, University of Michigan .....February 2, 2001
Paul Hovland, Argonne National Laboratory .....January 31, 2001
David Bailey, Lawrence Berkeley National Laboratory .....January 25, 2001
Robert Lucas, Lawrence Berkeley National Laboratory .....January 25, 2001
Gregory Miller, University of California, Davis .....January 19, 2001
Travis Austin, University of Colorado .....January 8-19, 2001
Steve McCormick, University of Colorado .....January 6-12, 2001
Stanimire Tomov, Texas A&M University .....January 3-12, 2001
Benjamin Keen, University of Michigan .....January 11, 2001
Eduardo D'Azevedo, Oak Ridge National Laboratory .....January 5, 2001
Esmond Ng, Lawrence Berkeley National Laboratory .....January 5, 2001
James Glimm, State University of New York, Stony Brook .....January 5, 2001
Erin Parker, University of North Carolina, Chapel Hill (Seminar) .....January 5, 2001
Mark Shephard, Rensselaer Polytechnic Research Center .....January 5, 2001
Calvin Lin, University of Texas, Austin (Seminar) .....January 4, 2001
Sally McKee, University of Utah .....December 13-15, 2000
Angela Shiflet, Wofford College .....December 14, 2000
Dana Knoll, Los Alamos National Laboratory .....December 13-14, 2000
Henry Tufo, Argonne National Laboratory .....December 4-14, 2000
Charles Breckenridge, SRC Computers, Inc. .....December 13, 2000
David Caliga, SRC Computers, Inc. .....December 13, 2000
Greg Fenner, SRC Computers, Inc. .....December 13, 2000
Michael Henesy, SRC Computers, Inc. .....December 13, 2000
Jon Huppenthal, SRC Computers, Inc. .....December 13, 2000
Daniel Poznanovic, SRC Computers, Inc. .....December 13, 2000
Ken Joy, University of California, Davis .....December 7, 2000
Robert Krasny, University of Michigan (Seminar) .....November 30-December 1, 2000
Tim Chartier, University of Colorado .....November 29-December 1, 2000
Ulrich Ruede, University of Erlangen (Seminar) .....November 2-30, 2000

Dawson Engler, Stanford University (Seminar) .....November 28, 2000
Matthew Gleeson, MPI Software Technologies Inc. .....November 20, 2000
Robert Sharpley, University of South Carolina .....November 17, 2000
Scott Johnson, University of South Carolina .....November 17, 2000
Peter Gottschling, GMD First (Seminar) .....November 16, 2000
Heinz Kreiss, University of California, Los Angeles .....November 13, 2000
Scott Gaffney, University of California, Irvine .....November 3, 2000
Padhraic Smyth, University of California, Irvine .....November 3, 2000
Christoph Pflaum, University of Wurzburg (Seminar).....October 24, 2000
Sanith Wijesinghe, Massachusetts Institute of Technology.....October 9-20, 2000
John Lyon, Dartmouth College .....October 17-20, 2000
Michael Wiltberger, Dartmouth College .....October 17-20, 2000
Raytcho Lazarov, Texas A & M University .....October 19, 2000
Stanimire Tomov, Texas A & M University .....October 19, 2000
Luis Silva, University of California, Los Angeles .....October 19, 2000
Ricardo Fonseca, University of California, Los Angeles (Seminar) ..... October 19, 2000
Warren Mori, University of California, Los Angeles .....October 19, 2000
Juan Alonso, Stanford University (Seminar) .....October 18, 2000
Antony Jameson, Stanford University (Seminar) .....October 18, 2000
William Bosl, Stanford University .....October 16, 2000
Erland Arge, Numerical Objects AS .....October 16, 2000
Are Magnus Bruaset, Numerical Objects AS (Seminar) .....October 16, 2000
Steve McCormick, University of Colorado .....October 4-6, 2000
Victor Barocas, University of Minnesota (Seminar) .....October 4-6, 2000
Klaus Stueben, GMD-Forschungszentrum Informationstechnik GmbH.....September 28-October 2, 2000

FY 2001 Institute for Terascale Simulation Lecture Series (in reverse chronological order)

Charles Peskin, Courant Institute of Mathematical Sciences (Seminar).....July 11, 2001
Eugene H. Spafford, Purdue University (Seminar) .....May 9, 2001
Michael Norman, University of California, San Diego (Seminar) .....March 28, 2001
Ingrid Daubechies, Princeton University (Seminar) .....January 24, 2001
David Johnson, AT&T (Seminar) .....November 15, 2000

Visiting Faculty, Guests, Consultants, and Researchers

Visiting and Collaborating Professors

Fernando Arias de Saavedra, University of Spain
Owe Axelsson, University of Niimegen
Randy Bank, University of California, San Diego
Martin Bertram, University of Utah
Marian Brezina, University of Colorado
Xiao-Chuan Cai, University of Colorado
Zhiqiang Cai, Purdue University

Visiting and Collaborating Professors (continued)

Alejandro Garcia, San Jose State University
Omar Ghattas, Carnegie Mellon University
Anne Greenbaum, University of Washington
Michael Holst, University of California, San Diego
Kenneth Joy, University of California, Davis
Johannes Kraus, Austrian Science Foundation
Raytcho Lazarov, Texas A&M University
Tara Madhyastha, University of California, Santa Cruz
Leszek Marcinkowski, University of Colorado
Sally McKee, University of Utah
Michael Minion, University of North Carolina
Joseph Pasciak, Texas A&M University
Francesco Pederiva, University of Trento, Italy
Christoph Pflaum, Technical University
John Ruge, Front Range Scientific Computing
Don Schwendeman, Rensselaer Polytechnic Institute
Klaus Stueben, GMD-Forschungszentrum Informations technik GmbH
Tonya Vassilevska, Bulgarian Academy of Sciences
Gabriel Wittum, Kiel University
Irad Yahneh, Technion Israel Institute of Technology
Jacob Ystrom, Royal Institute of Technology

Participating Guests

Fernando Arias de Saavedra, University of Spain
Marsha Berger, New York University
William Bosl, Stanford University
Marian Brezina, University of Colorado
George Byrne, Illinios Institute of Technology
Richard Cook, Univeristy of California, Davis
Roger Crawfis, Ohio State University
Eric de Sturler, University of Illinois
David Dean, University of Colorado
John Fitzgerald, Lawrence Livermore National Laboratory (retired)
Sharon Frazier, Lawrence Livermore National Laboratory (retired)
Kyle Gallivan, Florida State University
Alejandro Garcia, San Jose State University
Michael Gertz, University of California, Davis
Michael Griebel, University of Bonn
Bernd Hamann, University of California, Davis
Ulf Hannebutte, Intel Corportation
Kenneth Joy, University of California, Davis
Johannes Kraus, Univesity of Leoben
Raytcho Lazarov, Texas A&M University

Byung Lee, University of Vermont
Ida Lozares, Lawrence Livermore National Laboratory (retired)
Kwan-Liu Ma, University of California, Davis
Michael Minion, University of North Caralina
Frank Mueller, North Carolina State Univesity
Beth Ong, Lawrence Livermore National Laboratory (retired)
Joseph Pasciak, Texas A&M University
Michael Pernice, University of Utah
Elbridge Gerry Puckett, University of California, Davis
John Rice, University of California, Berkeley
Ulrich Ruede, University of Erlangen
Yousef Saad, University of Minnesota
Paul Saylor, University of Illinois
Daniel Schikore, Computational Engineering, International
Gregory Schussman, University of California, Davis
Rob van der Wijngaart, NASA Ames Research Center
Tonya Vassilevska, Bulgarian Academy of Sciences
Gabreil Wittum, Kiel University
Donald Wolitizer, California State University, Hayward
Jacob Ystrom, Royal Institute of Technology
Ludmil Zikatanov, Penn State University

Consultants

Bernie Alder, University of California (Professor Emeritus)
Randolph Bank, University of California, San Diego
Leo Breiman, University of California, Berkeley
Nancy Collins, University of Colorado, Boulder
Gene Golub, Stanford University
Anne Greenbaum, University of Washington
Charles Hansen, University of Utah
Michael Holst, University of California, San Diego
David Keyes, Old Dominion University
Heinz-Otto Kriess, University of California, Los Angeles
Luc Machiels, Swiss Federal Institute of Technology
Thomas Manteuffel, University of Colorado, Boulder
Stephen McCormick, University of Colorado, Boulder
Gregory Miller, University of California, Davis
Linda Petzold, Univerity of California, Santa Barbara
Steve Schaffer, New Mexico Tech
Homer Walker, Worcester Polytechnic Institute

Department of Applied Science Faculty

Nelson Max
Garry Rodrigue

Postdoctoral Researchers

Robert Anderson
Erick Cantu-Paz
Paul Castillo
Leonardo Colletti
Miguel Dumett
Petri Fast
Jean-Luc Fattebert
Jeff Hittinger
David Hysom
Bobby Philip
Markus Schordan
Leonid Tsap

University Collaborative Research Program Faculty and Students

Darrell Long and Zachary Peterson, University of California, Santa Cruz
B. S. Manjunath and Jelena Teslic, University of California, Santa Barbara
Linda Petzold and Yang Cao, University of California, Santa Barbara
Sutanu Sarkar, David Lopez, and Carlos Pantano, University of California, San Diego
Padhraic Smyth and Scott Gaffney, University of California, Irvine
Mark van der Laan and Annette Molinaro-Clark, University of California, Berkeley

LDRD Project Investigators

Bronis de Supinski, LLNL, Center for Applied Scientific Computing
Mark Duchaineau, LLNL, Center for Applied Scientific Computing
Chandrika Kamath, LLNL, Center for Applied Scientific Computing

Students

Student Guests

Nathan Crane, University of Illinois
Matt Giamporcaro, Boston University
Charles Hindman, University of Colorado
Jason Hunt, University of Michigan
McKay Hyde, California Institute of Technology
David Hysom, Old Dominion University
Ty Jones, Univerity of Nevada, Reno
Lars Karlsson, Chalmers University of Technology
Michael King, University of Utah
Justin Koo, University of Michigan, Ann Arbor

Falko Kuester, University of California, Davis
John Lai, University of California, Davis
Tushar Mohan, University of Utah
Sandra Naegele, University of Heidelberg
Diem Phuong Nguyen, University of Utah
Stefan Nilsson, Chalmers Institute of Techology
Christopher Oehmen, University of Tennessee
Erin Parker, University of North Carolina
Pete Poulos, University of Utah
Jonathan Rochez, University of California, Davis
Preethy Vaidyanath, University of California, Santa Cruz
Wing Yee, University of Utah

Department of Applied Science Students

Paul Covello
Ben Gregorski
Aaron Herrnstein
Ana Iontcheva
Joseph Koning
Daniel Laney
Tim Pierce
Jonathan Rochez
Robert Rieben
Bahrad Sokhansanj
Jay Thomas

ISCR Students

Cheryl Barkauskas, Washington University
Rita Borgo, University of Pisa, Italy
Timo Bremer, University of California, Davis
Oliver Broeker, Swiss Federal Institute of Technology
Gyu Sang Choi, Penn State University
Tom Dossa, Santa Clara University
Achim Gordner, University of Heidelberg
Chaz Hales, Brigham Young University
Rachel Knop, West Point
Tzanio Kolev, Texas A&M University
Markus Kowarschik, University of Erlangen
Stefan Lang, University of Heidelberg
Michael McCracken, Penn State University
Kathleen Metz, Las Positas College
Deanna Midtaune, University of Pacific
Moon Gyu Park, Purdue University
Serban Porumbescu, University of California, Davis
Joshua Senecal, University of California, Davis

ISCR Students (continued)

Stanimire Tomov, Texas A&M University
Nicolas Valette, Texas A&M University
Serge van Criekingen, Norhtwestern University
Sanith Wijesinghe, Massachusetts Institute of Technology
Yihao Zheng, University of California, Davis

ITST Students

Lucas Ackerman, Worcester Polytechnic
Ryan Avecilla, University of Illinois
Janine Bennett, University of California, Davis
David Buttler, Georgia Tech
Todd Coffey, North Carolina State University
Kree Cole-McLaughlin, University of Utah
Paul Dostert, Texas A&M University
Michael Flanagan, Texas A&M University
David Hysom, Old Dominion University
Diana Jackson, Wofford College
David Littau, University of Minnesota
Luke Olsen, University of Colorado
Min Shin, University of South Florida

National Physical Science Consortium (NPSC) Students

Rachel Karchin, University of California, Santa Cruz
Imelda Kirby, University of Washington
Megan Thomas, University of California, Berkeley

Workshops and Conferences

Bay Area Scientific Computing Day, LLNL, February 2001
Mining Scientific Datasets, Chicago, Illinois, April 2001
Copper Mountain Conference, Copper Mountain, CO, April 2001
Preconditioning 2001, Tahoe City, CA, April 2001
Linear Solvers Workshop, Livermore, CA, June 2001
Common Component Architecture (CCA) Workshop, Livermore, CA, July 2001
Sensitivity Workshop, Livermore, CA, August 2001

Institute for Scientific Computing Research

Seminar Series Abstracts

(in reverse chronological order)

September 28, 2001

QMTest and QMTrack: Testing Code and Tracking Progress

Mark Mitchell
CodeSourcery, LLC

Email: mark@codesourcery.com

Abstract:

The QMTest and QMTrack tools are the result of a Software Carpentry (www.software-carpentry.com) development project. These portable, open source tools use an extensible, programmable design to allow customization by users.

An overview of the design of both tools, as well as a demonstration of QMTest, will be provided. CodeSourcery's plans for the future of these tools will also be discussed.

The audience will be encouraged to provide suggestions about features that would be of particular utility in their development projects.

Research web page: http://www.software-carpentry.com/

Institution web page: http://www.codesourcery.com/

September 27, 2001

Simulation of Compressible Turbulent Flows With Reaction

Sutanu Sarkar
University of California, San Diego

Email: sarkar@mechanics.ucsd.edu

Abstract:

Two outstanding problems in the area of compressible reactive flows will be discussed. The first problem concerns large eddy simulation where the strongly nonlinear dependence of the reaction rate term on temperature leads to a subgrid contribution in the resolved-scale equations, which must be modeled. We have developed a subgrid model by using the information available in the resolved scales along with additional physical input, namely, a model spectrum for the unresolved scales. Promising results have been obtained when evaluating this subgrid reaction rate model against a direct numerical simulation of a shear layer. The second problem occurs when a burn initiated in a NIF capsule encounters inhomogeneities of mixture fraction due to the introduction of inert shell material into DT mix by Rayleigh–Taylor or Richtmeyer–Meshkov instabilities. Preliminary results have been obtained in a simple model problem to identify the important parameters that control the modified burn propagation.

Speaker's web page: http://www-mae.ucsd.edu/RESEARCH/SARKAR/sarkar.html

Institution web page: http://www.ucsd.edu/

September 25, 2001

The Behavior of Converging Shock Waves of Spherical and Polyhedral Form

Don Schwendeman

Rensselaer Polytechnic Institute

Email: schwed@rpi.edu

Abstract:

In this talk the behavior of converging spherical shock waves will be discussed. An analysis of converging shocks, whose initial shape takes the form of regular polyhedra, will be presented within the framework of Whitham's theory of geometrical shock dynamics. The analysis of this problem is motivated by the earlier work on converging cylindrical shocks discussed in Schwendeman and Whitham (*Proc. R. Soc. Lond.*, A413, 297-311, 1987). In that paper, exact solutions were reported for converging polygonal shocks in which the initial shape reforms repeatedly as the shock contracts. For the polyhedral case, the analysis is performed both analytically and numerically for an equivalent problem involving shock propagation in a converging channel with triangular cross section.

It is found that a repeating sequence of shock surfaces composed of nearly planar pieces develops, although the initial planar surface does not reform, and that the increase in strength of the shock at each iterate in the sequence follows the same behavior as that for a converging spherical shock independent of the convergence angle of the channel. In this sense, the shocks are stable and the result is analogous to that found in the two-dimensional case discussed in the earlier paper. A numerical study of converging spherical shocks subject to smooth initial perturbations in strength shows a strong tendency to form surfaces composed of nearly planar pieces, suggesting that the stability result is fairly general.

Institution web page: http://www.rpi.edu/

September 14, 2001

Multilevel Conjugate Gradient Methods for Nonlinear Least-Squares Finite Element Computations

Gerhard Starke
University of Hannover

Email: starke@ifam.uni-hannover.de

Abstract:

Least-squares, mixed-finite element discretizations of nonlinear elliptic boundary value problems lead to nonlinear algebraic least-squares problems of large dimension. This talk describes the use of nonlinear conjugate gradient methods for these problems. The gradient of the functional and, consequently, the search directions are computed with respect to the natural norm of the underlying variational problem. This is crucial in order to achieve convergence rates that are independent of the refinement level of the triangulation. An appropriate multilevel inner product on the space where the variational problem is formulated allows for the efficient implementation of the method. We conclude with computational results for a sequence of discretizations based on adaptively refined triangulations for a variably saturated subsurface flow problem.

Speaker's web page: http://www.ifam.uni-hannover.de/~starke/

Institution web page: http://www.uni-hannover.de/

September 6, 2001

Introduction to Architecture and Performance of First European TeraFlops Supercomputer: The Hitachi SR8000

Gerhard Wellein
University of Erlangen-Nuernberg

Email: gerhard.wellein@rrze.uni-erlangen.de

Abstract:

Last year, the LRZ Munich released Europe's first TeraFlops computer for public user service, a Hitachi SR8000-F1. The Hitachi SR8000 series, which is based on RISC technology, was designed to incorporate the advantages of vector and scalar/RISC computing. This talk introduces the concept of pseudo-vector processing at the processor level as well as the flexible usage model, which supports both MPP-like and vector/SMP-like programming style. Performance characteristics for basic kernel loops and complex large-scale application codes will be discussed, including comparisons with other present-day MPP and vector supercomputers.

Research web page: http://www.hlrs.de/

September 6, 2001

Input/Output Scalability of Different Architectures

Preethy Vaidyanathan
University of California, Santa Cruz

Email: vaidyanathan1@llnl.gov

Abstract:

We characterize the I/O behavior of a computational biology application on Linux clusters with different file systems and on an IBM SP/2. This application plays a central role in the Human Genome Project. We show that locality is a very important factor affecting the performance of this application. We present the design of a user-level library for a new model of location-transparent storage to automatically redirect read accesses to the most appropriate location.

Institution web page: http://www.ucsc.edu/

August 31, 2001

Prefetching and Caching Strategies for Modern Memory Systems

Steven Reinhardt
University of Michigan

Email: stever@eecs.umich.edu

Abstract:

Microprocessor technology advances continue to provide larger on-chip caches and higher memory bandwidths, but off-chip memory latency is and will remain a significant bottleneck for system performance. Our research concerns how best to exploit growing capacity and bandwidth to counter the impact of latency.

We seek to exploit memory bandwidth by improving the effectiveness of prefetching. Traditional prefetching schemes risk degrading performance unless accuracies are high. Our prefetching framework effectively eliminates bandwidth contention by prioritizing requests, and reduces pollution effects via cache placement. Using this framework, aggressive prefetching of large memory regions boosts performance dramatically when applications have spatial locality.

We seek to exploit the capacity of on-chip caches more fully by managing their contents intelligently. Virtual-memory systems use full associativity and software-controlled replacement to combat the high latency of disk transfers; we are exploring the extent to which these ideas can be applied to large on-chip caches. We have developed a practical fully associative cache structure and a novel software replacement algorithm tailored for secondary caches. In tandem, they can achieve significant miss-rate reductions compared to conventional organizations.

Future work includes investigating the potential of application and operating-system involvement in secondary-cache management.

Speaker's web page: http://www.eecs.umich.edu/~stever/

Institution web page: http://www.umich.edu/

August 28, 2001

A Comparison of Factorization-free Eigensolvers with Application to Cavity Resonators

Peter Arbenz
Swiss Federal Institute of Technology

Email: arbenz@inf.ethz.ch

Abstract:

We investigate eigensolvers for the generalized eigenvalue problem $Ax = \lambda Mx$ with symmetric A and symmetric positive definite M that do not require the factorization of either A or M. We compare various variants of preconditioned Rayleigh quotient minimization and preconditioned Jacobi–Davidson algorithm by means of large-scale finite element discretizations originating from the design of the new RF cavity ring cyclotron installed at the Paul Scherrer Institute (PSI) in Villigen, Switzerland.

Institution web page: http://www.epfl.ch/

August 24, 2001

Interactive Visualization of Large Graphs and Networks

Tamara Munzner
Compaq Systems Research Center

Email: tamara.munzner@compaq.com

Abstract:

Many real-world domains can be represented as large node-link graphs: backbone Internet routers connect with 70,000 other hosts, mid-sized Web servers handle between 20,000 and 200,000 hyperlinked documents, and dictionaries contain millions of words defined in terms of each other. Computational manipulation of such large graphs is common, but previous tools for graph visualization have been limited to datasets of a few thousand nodes.

Visual depictions of graphs and networks are external representations that exploit human visual processing to reduce the cognitive load of many tasks that require understanding of global or local structure. We assert that the two key advantages of computer-based systems for information visualization over traditional paper-based visual exposition are interactivity and scalability. We also argue that designing visualization software by taking the characteristics of a target user's task domain into account leads to systems that are more effective and scale to larger datasets than previous work.

This talk contains a detailed analysis of three specialized systems for the interactive exploration of large graphs, relating the intended tasks to the spatial layout and visual encoding choices. We present two novel algorithms for specialized layout and drawing that use quite different visual metaphors. The H3 system for visualizing the hyperlink structures of Web sites scales to datasets of over 100,000 nodes by using a carefully chosen spanning tree as the layout backbone, 3D hyperbolic geometry for a Focus+Context view, and provides a fluid interactive experience through guaranteed frame rate drawing. The Constellation system features a highly specialized 2D layout intended to spatially encode domain-specific information for computational linguists checking the plausibility of a large semantic network created from dictionaries. The Planet Multicast system for displaying the tunnel topology of the Internet's multicast backbone provides a literal 3D geographic layout of arcs on a globe to help MBone maintainers find misconfigured long-distance tunnels.

Each of these three systems provides a very different view of the graph structure, and we evaluate their efficacy for the intended task. We generalize these findings in our analysis of the importance of interactivity and specialization for graph visualization systems that are effective and scalable.

Speaker's web page: http://graphics.stanford.edu/~munzner/

Institution web page: http://www.compaq.com/

August 22, 2001

3D Pore Structure/Fluid Distribution Measurement and Network Model Flow Studies

Brent Lindquist

State University of New York, Stony Brook

Email: lindquis@ams.sunysb.edu

Abstract:

X-ray computed microtomography (XCMT) has made it possible to create three-dimensional images of the pore space of rock at resolutions around 5 microns. A 1024^2 digitized image thus "captures" a cube of rock with a measurement of one-half cm on a side. The availability of such images has revitalized efforts to understand the connection between pore microstructure and core scale rock flow properties.

Our contribution to this effort has been the development of image analysis algorithms to extract 3D stochastic descriptions of pore geometry parameters. We present results for an analysis of a suite of Fontainbleau sandstones from 7.5 percent to 22 percent porosity. Parameters measured include distributions of pore coordination number, channel length, throat size, and pore volume, as well as nearest neighbor correlation distributions for throat-throat and throat-pore sizes.

This experimental analysis provides perfect input for network model flow simulations. The results obtained (average coordination numbers around 3.5, log-normal pore-size distributions, exponentially tailed throat size distributions) have profound impact on network models, which, in the absence of precise experimental measurements, have resorted to idealized geometrical networks. We present results performed in conjunction with our collaborators at Australian National University, demonstrating the profound effects that the realistic network models have on simulation results for residual oil saturations in invasion-percolation based models for two-phase flow.

Experimental techniques to image fluids in XCMT are beginning to improve. Because XCMT is non-invasive, time series data is possible. We discuss preliminary results on the intra-pore distributions of water and oil as a result of both water- and oil-driven hysteretic flows.

Speaker's web page: http://www.ams.sunysb.edu/~lindquis/lindquist.html

Institution web page: http://www.sunysb.edu/

August 21, 2001

Multidimensional Upwinding, Preconditioning, and Other Stuff from Michigan

Philip Roe
University of Michigan

Email: philroe@engin.umich.edu

Abstract:

This is a brief synopsis of current research in the University of Michigan Aerospace Engineering Department related to computational fluid dynamics. The faculty includes Bram van Leer, Ken Powell, Iain Boyd, and me. Topics include magnetohydrodynamics, rarified flow, elastic waves, and high-order schemes.

Methods that incorporate PDE information at a higher-than-one-dimensional level are discussed. Potential or realized benefits include:

1. Valid incompressible limits of compressible codes
2. Faster convergence to steady state
3. Solution-driven grids
4. Economically captured discontinuities
5. Preservation of div/curl constraints
6. Avoidance of computational anomalies

Institution web page: http://www.umich.edu/

August 13, 2001

Computational Study of Arterial Flow Disturbance Induced by Intravascular Stents

Levanto Schachter
University of California, Davis

Abstract:

A common interventional protocol against advanced cardiovascular disease involves the placement of an intravascular stent, an expandable wire mesh structure, into the diseased artery. The primary limitation of stenting procedures is restenosis, a complex process by which vessel blockage redevelops over a period of a few months. The deployment of a stent within an arterial segment locally damages the endothelium, the monolayer of cells lining the inner surfaces of blood vessels. As a result of this damage, both thrombotic and wound healing biological pathways are rapidly activated.

It appears that the incidence of restenosis is reduced if re-endothelialization of the damaged vascular area occurs sufficiently rapidly. Re-endothelialization requires endothelial cell proliferation and migration, and these processes are likely affected by the local hemodynamic environment. We have been using computational fluid dynamics (CFD) to study the detailed flow field in the vicinity of implanted stents. Of particular interest is the extent of flow disturbance induced by the placement of a stent within an artery. The simulations involve solving the full three-dimensional time-dependent equations governing fluid mass and momentum conservation (Navier–Stokes equations) in model straight and curved arterial segments. The effects of various parameters including stent wire thickness, inter-wire spacing, extent of vessel curvature, flow Reynolds number, and flow inlet inlet velocity profile on the occurrence of flow disturbance have been investigated.

Our results have revealed a region of flow separation and recirculation immediately downstream of model stents. The size of this recirculation flow zone increases with both the flow Reynolds number and the stent wire thickness. With flow pulsatility, this region of flow disturbance periodically appears and disappears. Interestingly, the nature of the flow disturbance downstream of a stent depends on a complex fashion on the vessel geometry, flow conditions, and stent design.

Given the wide spectrum of arterial configurations and flow conditions present in vivo, an interesting notion would be to use CFD techniques to optimize the design of stents for particular vascular sites with the goal of minimizing stent-induced flow disturbance.

August 9, 2001

New Monte Carlo Methods for Problems in Materials and Biology

Michael Mascagni
Florida State University

Email: mascagni@cs.fsu.edu

Abstract:

Probabilistic potential theory enables us to solve a large class of parabolic and elliptic partial differential equations using diffusion techniques. Here, we present new first- and last-passage Monte Carlo algorithms and show their utility in materials science and biology problems. These techniques exploit the fact that the first-passage probability function is the Green's function for the Dirchlet problem of the Laplace equation. First-passage algorithms allow the rapid simulation of diffusion using analytic or simulation-based Green's functions in rather less complicated basic geometries. This permits consideration of more complicated real geometries made up as combinations of the simple geometries where Green's functions are available. This new method is the extension of the well-known "Walk on Spheres" method.

Harnessing these first-passage algorithms, we have developed the fastest algorithms known to compute (1) the fluid permeability in overlapping, nonoverlapping, and polydispersed spherical models of random porous media, (2) the Solc-Stockmayer model with zero potential, a model of ligand binding, (3) the mean trapping rate of a diffusing particle in a domain of nonoverlapping spherical traps, and (4) the effective conductivity for perfectly insulating, nonoverlapping spherical inclusions in a matrix of conductivity.

In certain problems, such as that of computing the electrostatic charge distribution on a conductor, using the last-passage distribution is useful. Using these analogous last-passage algorithms, we have solved the test problem of computing the charge distribution on a circular two-dimensional disk in three dimensions.

Our plans for the future involve adding more surface Green's functions to our present set of known Green's functions, and the application of these techniques to more realistic problems in materials and biology.

This is joint work with Dr. Chi-Ok Hwang of Florida State University, and Dr. James Given of Angle, Inc.

Speaker's web page: http://www.cs.fsu.edu/~mascagni/

Institution web page: http://www.fsu.edu/

August 8, 2001

Effective Query Processing Over Large Multidimensional Data Sets

Sharad Mehrotra
University of California, Irvine

Email: sharad@ics.uci.edu

Abstract:

In this talk, I will present two of our recent research results in query processing over multidimensional data sets. The first result deals with answering aggregate queries in an approximate manner over very large data collections. Our approach uses selective traversal of a multiresolution aggregate (MRA) tree structure to provide 100% interval of confidence on the value of the aggregate. The algorithm works iteratively coming up with improving quality answers until some error metric is satisfied or time constraint is reached. This work is part of the Quality Aware Sensor Database Architecture (QUASAR) project, in which we explore techniques to trade quality of answer for performance in database query processing when in highly dynamic data intensive environments.

The second part of the talk will discuss local dimensionality reduction (LDR) approach that effectively overcomes the dimensionality curse problem in range and K-NN searches over highly multidimensional data sets. Unlike previous approaches, LDR exploits the local correlation in data to identify clusters that have a natural lower-dimensional representation. These clusters can then be individually indexed and searched using multidimensional data structures. This work is part of the Multimedia Analysis and Retrieval Systems (MARS) project in which we are developing a data management system that provides native and scalable support for similarity retrieval and query refinement over multimedia information.

Speaker's web page: http://www-db.ics.uci.edu/pages/people/sharad.shtml

Research web page: http://www-db.ics.uci.edu/pages/research/mars/

Institution web page: http://www.uci.edu/

July 27, 2001

Performance Analysis of MPI/OpenMP Applications Using Paraver

Jesus Labarta

European Center for Parallelism of Barcelona

Email: jesus@cepba.upc.es

Abstract:

Paraver is a visualizer for the performance of parallel programs that supports detailed analyses down to very fine granularities. Two major strengths of Paraver are:

- Flexibility and support for analyses not even planned.
- Accurate quantitative analysis support.

The OMPItrace instrumentation package for MPI, OpenMP, and mixed mode programs has been ported to IBM SP platforms in cooperation with the ACTC. Besides the process and communication activity, a wide range of performance indices derived from hardware counters can be displayed and measured. A translator from the UTE trace format to Paraver is also available.

Dimemas is a simulator useful in performance estimation of message-passing programs. Dimemas rebuilds the timing behavior of MPI programs based on simple target architecture parameters. These parameters describe basic aspects that influence application performance such as network latency and bandwidth, injection mechanism, processor performance, and predicted effect of improvement in one code section.

The presentation will give an update of the type of analyses that can be performed with Paraver and Dimemas as well as their interoperability capabilities. Examples will be mostly based on the Sweep3d application. For more information see: http://www.cepba.upc.es/paraver/ and http://www.cepba.upc.es/dimemas/.

Institution web page: http://www.cepba.upc.es/

July 26, 2001

Performance Optimization of Component-based Data Intensive Applications

Alan Sussman
University of Maryland

Email: als@cs.umd.edu

Abstract:

In the DataCutter project we have proposed the use of a component-based processing model, called filter-stream programming, for data intensive applications. In this talk I will give an overview of filter-stream programming in DataCutter, and also discuss issues related to retrieval of data stored in a distributed, high-performance, heterogeneous computing environment. The main part of the talk will discuss several of the optimization techniques we are designing and evaluating on the heterogeneous set of computational and storage Linux clusters at the University of Maryland. These optimizations include the reuse of existing groups of filters to perform a requested computation, and the transparent replication of particular filters that cause bottlenecks in a filter pipeline. I will present results from the evaluation showing that performance varies greatly depending on the assignment of filters to hosts and the number of transparent copies instantiated for each filter, and that performance is also highly dependent on the distribution of source data to host disks. In ongoing work, we are using these results to drive the construction of cost models to enable automating the placement of filters on the set of available hosts, and selecting the number of copies of each filter on each host.

Speaker's web page: http://www.cs.umd.edu/~als/

Institution web page: http://www.umd.edu/

July 25, 2001

A Multilevel Nonlinear Method

Irad Yavneh

Technion-Israel Institute of Technology

Email: irad@cs.technion.ac.il

Abstract:

Robust multigrid algorithms for linear boundary-value problems are well researched. For nonlinear problems, two general approaches are used widely: Global Linearization (GL) and Local Linearization (LL). In GL methods, the discretized problem is linearized, usually by Newton's method, and the resulting linear system is solved approximately by a linear multigrid algorithm. This is repeated iteratively. In the LL approach, the nonlinear fine-grid operator is approximated nonlinearly on the coarser grids, and explicit linearization is only performed locally, in the relaxation process. The best known of the LL methods is the so-called Full Approximation Scheme (FAS). For simple problems, the two approaches often perform similarly, but a distinct behavior is exhibited in more complicated settings, with the GL approach performing better in some cases and the LL approach in others.

We propose a Multilevel Nonlinear Method (MNM) which, we believe, will generally be at least as robust as either one of the above, and often more robust than both. The work described is at an early stage, and only a highly simplified analysis and preliminary numerical results are yet available. However, these indicate that MNM may continue to perform even as the FAS and Newton methods melt in your hand.

Institution web page: http://www.technion.ac.il/

July 24, 2001

Three-Dimensional Control-Volume Mixed Finite Element Methods and Efficient Algebraic Solvers for Distorted Hexahedral Grids

Thomas Russell
University of Colorado, Denver

Email: trussell@carbon.cudenver.edu

Abstract:

The control-volume mixed finite element (CVMFE) method is designed to calculate accurate pressure and velocity distributions in subsurface flow problems with irregular geology and heterogeneous, discontinuous, anisotropic conductivity. In two dimensions, it differs from the usual lowest-order Raviart–Thomas (RT) mixed method in its choice of velocity-vector test functions, which lead to discrete analogues of Darcy's law on control volumes. Irregular geology is modeled with logically rectangular distorted quadrilateral and hexahedral grids in 2D and 3D, respectively.

The extension from 2D to 3D raises surprising issues and dilemmas. For example, in general, the trial space obtained from the standard Piola transformation does not contain the uniform constant velocity field. The 3D formulation is presented with details about better choices of trial functions, along with a sampling of numerical results that show second-order convergence when the exact solution is not singular. Also presented, in the context of the RT method with a forthcoming extension to CVMFE, is an efficient solver for the 3D equations. This uses a convenient basis for the divergence-free velocity functions to reduce the equations to a symmetric positive-definite system of smaller size than has been possible previously. With conjugate gradients preconditioned by overlapping domain decomposition, this enables the 3D equations to be solved with effort comparable to that for other methods.

Institution web page: http://www.cudenver.edu/

July 19, 2001

Two Performance Studies in Support of the NSF/DOE Coupled Climate Model

Patrick Worley
Oak Ridge National Laboratory

Email: worleyph@ornl.gov

Abstract:

We describe two performance studies that were undertaken in support of the development and optimization of the NSF/DOE Coupled Climate Model. The first study motivates the design of a new data structure for the column physics routines used in the atmospheric model. The new data structure allows index ordering and loop ordering and length to be set at compile or runtime. We use empirical studies on Compaq AlphaServer SC, IBM SP, and NEC SX5 systems to demonstrate the performance gains that are made possible by this type of flexibility.

The second study is an evaluation of the "allreduce" collective communication operator on the Compaq AlphaServer SC and IBM SP. We compare the performance of the vendor-supplied MPI_ALLREDUCE with that of a number of different algorithms for a variety of vector lengths, processor counts, and assumptions on how the allreduce is used in an application code. Best observed performance is used to identify good parallel algorithms and to compare performance between systems.

Speaker's web page: http://www.csm.ornl.gov/~worley/

Research web page: http://www.csm.ornl.gov/evaluation/

Institution web page: http://www.ornl.gov/

July 18, 2001

A Fast Poisson Solver for Irregular 3D Regions

Anne Greenbaum
University of Washington

Email:
greenbau@math.washington.edu

Abstract:

We describe a fast method for solving Poisson's equation on irregular 3D regions. The method consists of two steps, each carried out with a fast Poisson solver (or a multigrid solver) applied on an embedding cube.

The first step is to evaluate a volume integral in order to obtain some solution to the inhomogeneous equation. Using known jumps in the second derivatives of the volume integral, one can determine the right-hand side for a discrete version of Poisson's equation on a uniform grid throughout an embedding cube. This equation can then be solved with an FFT or a multigrid method. Next, one solves a boundary integral equation to determine an integral representation of the solution to Laplace's equation with appropriate boundary conditions. Knowing this integral representation, one can compute the discontinuities in the solution and its derivatives across the boundary of the region and use these to determine the right-hand side for another fast Poisson solve on the embedding cube. The total cost is that of two fast Poisson solves on a cube [O(N log N) or O(N), depending on whether an FFT or a multigrid method is used] plus the cost of solving a boundary integral equation [O(m^2) or O(m), where m is the number of boundary discretization points, depending on whether the dense *m* by *m* matrix is assembled or the fast multipole method is used to compute matrix-vector products].

A software package implementing these ideas is currently being constructed. The algorithms are amenable to parallelization, and specific problems have been solved using multiple processors on an IBM SP-2.

Speaker's web page: http://www.washington.edu/~greenbau/

Institution web page: http://www.washington.edu/

July 18, 2001

Parallel Data Access in Heterogeneous Environments with Storage Tank

Randal Burns
IBM Almaden Research Center

Email: randal@almaden.ibm.com

Abstract:

We describe how the Storage Tank distributed file system manages platform and network heterogeneity in a large-scale computing environment. Storage Tank provides shared access from multiple operating systems to fibre channel, iSCSI, or object-based network attached storage. This allows multiple computers or clusters to share data without degrading performance at compute nodes.

We also propose data placement algorithms that de-cluster I/O and place data near where they are used in heterogeneous networks. Declustering will help Storage Tank achieve aggregate write rates much greater than the capacity of any single switch or link in the system. Storage Tank declusters data on a per-file basis using a policy engine, and within a single file through distributed storage pools, both of which adapt to changing workloads and component failures.

Speaker's web page: http://www.cse.ucsc.edu/~randal/

Institution web page: http://www.almaden.ibm.com/

July 17, 2001

TAU Performance System: Developments and Evolution

Allen Malony
University of Oregon

Email: malony@cs.uoregon.edu

Abstract:

The TAU performance system offers a robust and portable performance instrumentation, measurement, and analysis framework and toolkit for its implementation. TAU is based on a general parallel computation model and is targeted to large-scale, multi-thread shared memory and distributed memory message-passing parallelism. With over eight years of research and development, the TAU performance system is mature and is now being applied in several ASCI environments.

This talk gives some highlights of current TAU developments and application. New tools include a new Java-based parallel profile display tool and a new Fortran90 front-end for our Program Database Toolkit. We will present case studies of TAU's use for multi-threaded and mixed-mode performance measurement, and also our work with the Utah Uintah parallelism framework being applied in the C-SAFE ASCI effort. This latter effort includes an initial version of an experimental management system that will be described. Recently, we have proposed and prototyped an OpenMP performance API. Its design and implementation will be presented in the talk.

The TAU performance system will evolve in several important directions that are a natural extension of current capabilities as well as integration with other key technologies. We will describe these future activities.

Speaker's web page: http://www.cs.uoregon.edu/~malony/

Research web page: http://www.cs.uoregon.edu/research/paracomp/tau/

Institution web page: http://www.uoregon.edu/

July 16, 2001

Parallel and Robust Multigrid Techniques on Structured Grids

Ignacio Llorente
Complutense University, Madrid

Email: llorente@eucmax.sim.ucm.es

Abstract:

Research presented at this seminar, carried out by DACYA/UCM (Departamento de Arquitectura de Computadores y Automatica/Universidad Complutense de Madrid) in collaboration with ICASE (Institute of Computer Applications in Science and Engineering), focuses on the development of parallel and robust multigrid techniques on structured grids. We present results from the parallel implementation of different robust alternatives to solve the 3D diffusion-convection operator in highly non-aligned anisotropic problems. This study was used as a starting point for the development of a robust solver for the 3D Navier–Stokes equations. The results include both algorithmic and architectural properties (convergence rate, efficiency, scalability, and cache-oriented issues) on two different architectures, namely a low-cost distributed memory architecture (CORAL Beowulf-class system) and a cc-NUMA shared memory machine (SGI Origin 2000).

Speaker's web page: http://www.dacya.ucm.es/nacho/

Research web page: http://www.dacya.ucm.es/nacho/research/research.html

July 13, 2001

Hybrid Intelligent Systems for Industrial Data Analysis

Arthur Kordon
The Dow Chemical Company

Email: akordon@dow.com

Abstract:

Industry has recently significantly increased its efforts toward effective data analysis. Many companies even look at industrial data analysis as one of the main driving forces for establishing strategic competitive advantage. Hybrid intelligent systems, which integrate various techniques like neural networks, genetic programming, support vector machines, fuzzy logic, knowledge acquisition, and so forth, can play a significant role in this process. The objective of the presentation is to summarize the research topics and implementation issues of hybrid intelligent systems for industrial data analysis.

The first part of the presentation is focused on the special features and issues of industrial data analysis. In contrast to research-driven data analysis where the objective is to "transfer data into knowledge," the goal of industrial-driven analysis is to "transfer data into value." The difference in the objective function has an inevitable effect on the whole strategy for data analysis. Of special importance is the constraint on the level of complexity of delivered models (i.e., there is no need to increase complexity if the economic objective is accomplished). This strategy pushes industrial data analysis toward simple and robust empirical solutions that can be achieved with minimal time and training efforts.

The following special features of industrial data analysis are discussed briefly: hidden effect of operators' intervention, "curse of closed loops" effect, multiple time scale data analysis, and real-time pressure. The current state of the art and some issues of industrial data analysis, such as non-adaptive nature of the delivered empirical models and lack of flexible integration of the most appropriate data analysis methods and knowledge sources, are covered.

The second part of the presentation is devoted to the use of hybrid intelligent systems in analyzing industrial data. The "classic" hybrid intelligent systems of neural nets, fuzzy logic, genetic algorithms, and rule-based knowledge are compared with the "modern" integration of stacked analytic neural nets (not based on back-propagation), support vector machines, and genetic programming. The open research topics for development of modern hybrid intelligent systems (e.g., kernel selection for support vector machines, VC-dimension for analytic neural nets, integrating knowledge in genetic programming, etc.), as well as the main application areas like adaptive soft sensors, real-time design of experiments, and data-driven fundamental model emulators are discussed. The potential of hybrid intelligent systems for industrial data analysis is shown with several applications in The Dow Chemical Company (robust soft sensor in a chemical reactor, automating operating discipline in a large-scale chemical plant, using genetic programming for minimizing the design of experiments, and data-driven emulators for complex fundamental models).

The final part of the presentation defines the main industrial needs for improving data analysis. On the top of the list is the need to integrate technical and business-related data analysis. This is of special importance to global companies operating in the real-time financial environment of the global market. Another key factor for successful mass scale implementation of this approach is the development of autonomous process data analysis in real time.

Institution web page: http://www.dow.com/

July 12, 2001

Implementing Parallel Shortest Path for Parallel Transportation Applications

Michelle Hribar
Pacific University

Email: hribarm@pacificu.edu

Abstract:

Shortest path algorithms are used by many different applications; network routing, VLSI design, and transportation algorithms are just a few. Shortest path computation can be very expensive. Algorithms can take up to *O(n**2)*, where *n* is the size of the network (number of nodes or vertices). As the size of the network grows, this shortest path computation can dominate the execution time of the application. Therefore, parallel processing can provide the compute power required to quickly solve large problems.

Parallel shortest path algorithms are a well-studied topic, particularly in the area of theoretical development. In this talk, I will discuss an experimental study of parallel shortest path algorithms that focuses on understanding the performance impact of different implementation issues. In particular, I will focus on three implementation areas: (1) choice of shortest path algorithm, (2) termination detection, and (3) network decomposition. I will show how implementation decisions in all three areas have a large impact on the communication and convergence of parallel shortest path algorithms. In general, we find that communicating the most information at a time results in the best convergence and overall execution time. This is contrary to most scientific applications where it is optimal to minimize communication. I will present these results in the context of a transportation application that solves for traffic equilibrium.

Speaker's web page: http://zeus.cs.pacificu.edu/~hribarm/

Institution web page: http://www.pacificu.edu/

July 9, 2001

Synthesizing Sounds from Physically Based Motion

James O'Brien
University of California, Berkeley

Email: job@cs.berkeley.edu

Abstract:

This talk describes a technique for approximating sounds that are generated by the motions of solid objects. The technique builds on previous work in the field of physically based animation that uses deformable models to simulate the behavior of the solid objects. As the motions of the objects are computed, their surfaces are analyzed to determine how the motion will induce acoustic pressure waves in the surrounding medium. The technique computes the propagation of those waves to the listener and then uses the results to generate sounds corresponding to the behavior of the simulated objects.

Speaker's web page: http://www.cs.berkeley.edu/~job/

Institution web page: http://www.berkeley.edu/

July 6, 2001

A Scalable Hierarchical Algorithm for Unsupervised Clustering

Daniel Boley
University of Minnesota

Email: boley@cs.umn.edu

Abstract:

Top-down hierarchical clustering can be done in a scalable way. Here we describe a scalable, unsupervised clustering algorithm designed for large datasets from a variety of applications. The method constructs a tree of nested clusters top-down, where each cluster in the tree is split according to the leading principal direction. We use a fast principal direction solver to achieve a fast overall method. The algorithm can be applied to any dataset whose entries can be embedded in a high-dimensional Euclidean space, and takes full advantage of any sparsity present in the data. We show the performance of the method on text document data, in terms of both scalability and quality of clusters. We demonstrate the versatility of the method in different domains by showing results from text documents, human cancer gene expression data, and astrophysical data. For that last domain, we use an out-of-core variant of the underlying method, which is capable of efficiently clustering large datasets using only a relatively small memory partition.

Speaker's web page: http://www.cs.umn.edu/~boley/

Institution web page: http://www.umn.edu/

July 6, 2001

Dynamic Meshes, Dynamic Interfaces, and Hemodynamics

Omar Ghattas
Carnegie Mellon University

Email: oghattas@cs.cmu.edu

Abstract:

Many important phenomena in science and engineering, including our motivating problem of microstructural blood flow, can be modeled as flows with dynamic interfaces. The major challenge faced in simulating such flows is resolving the interfacial motion. Lagrangian methods are ideally suited for such problems, since interfaces are naturally represented and propagated. However, the material description of motion results in dynamic meshes, which become hopelessly distorted unless they are regularly regenerated. Lagrangian methods are particularly challenging on parallel computers, because scalable dynamic mesh methods remain elusive.

I will present a parallel dynamic mesh Lagrangian method for flows with dynamic interfaces that we have been developing at CMU. We take an aggressive approach to dynamic meshing by triangulating the propagating grid points at every time step using a scalable parallel Delaunay algorithm. Contrary to conventional wisdom, I will provide evidence that the costs of the geometric components (triangulation, coarsening, refinement, and partitioning) can be made small, relative to the flow solver. For example, in a 2D simulation of 10 interacting viscous cells with 500,000 unknowns on 64 processors of a Cray T3E, dynamic meshing consumes less than 5% of a time step. Moreover, our experiments on up to 128 processors show that the computational geometry scales about as well as the flow solver.

I will discuss the application of our dynamic mesh Lagrangian method to microstructural simulation of blood flow, which is essentially a problem in modeling the interaction of fluid-solid mixtures. The model is termed "microstructural" because it distinguishes the fluid (blood plasma and hemoglobin) from the solid (cell membrane) at micron scales and computes the momentum exchange between them. This is in contradistinction to typical macroscopic models that treat blood as a homogeneous viscous medium with phenomenological incorporation of cellular effects. I will conclude with a discussion of the prospects for microstructural modeling of blood flow at scales of interest in the design of artificial heart devices.

This work is jointly conducted with graduate student Ivan Malcevic, CMU colleagues Guy Blelloch, Gary Miller, and Noel Walkington, and University of Pittsburgh Medical Center collaborator Jim Antaki.

Speaker's web page: http://www.cs.cmu.edu/~oghattas/omar.html

Institution web page: http://www.cmu.edu/

June 25, 2001

Relaxation Schemes and Approximate Riemann Solvers for Nonlinear Hyperbolic Problems with Variable Coefficients

Randall LeVeque
University of Washington

Email: rjl@amath.washington.edu

Abstract:

A nonlinear system of *m* conservation laws can be rewritten as a linear system of *2m* equations, with a stiff source term that drives the solution towards that of the original equation. Relaxation schemes, numerical methods based on this idea, were introduced by Jin and Xin and have been used for many problems. These methods can be reinterpreted as a way to define an approximate Riemann solver for the original problem. This viewpoint suggests a broader class of "relaxation Riemann solvers" that approximate the solution to a system of *m* equations using *2m* waves. Nonlinear conservation laws with spatially varying flux functions can often be solved using an approximate Riemann solver of this form. One interesting application is to nonlinear elastodynamics in a laminate of thin layers, and hence rapidly varying material parameters. An initially smooth disturbance starts to sharpen into a shock but becomes oscillatory and appears to break up into a train of solitons due to the dispersive nature of the heterogeneous material.

Speaker's web page: http://www.amath.washington.edu/~rjl/

Research web page: http://www.amath.washington.edu/~claw/

Institution web page: http://www.washington.edu/

June 19, 2001

Efficient Adaptive Simplification of Massive Meshes for Visualization

Eric Shaffer

University of Illinois, Urbana-Champaign

Email: shaffer1@cs.uiuc.edu

Abstract:

The growing availability of massive polygonal models, and the inability of most existing visualization tools to work with such data, has created a pressing need for memory efficient methods capable of simplifying very large meshes. In this talk, we will present a method for performing adaptive simplification of polygonal meshes that are too large to fit in-core.

Our algorithm performs two passes over an input mesh. In the first pass, the model is quantized using a uniform grid, and surface information is accumulated. This sampling is then used to construct a BSP-Tree in which the partitioning planes are determined by the surface information. In the final pass, the original vertices are clustered using the BSP-Tree, yielding an adaptive approximation of the original mesh. The BSP-Tree describes a natural simplification hierarchy, making it possible to generate a progressive transmission and construct level-of-detail representations. In this way, the algorithm provides some of the features associated with more expensive edge contraction methods while maintaining greater computational efficiency. In addition to performing adaptive simplification, our algorithm exhibits output-sensitive memory requirements and allows fine control over the size of the simplified mesh.

Speaker's web page: http://bugle.cs.uiuc.edu/People/shaffer1/

Institution web page: http://www.uiuc.edu/

June 15, 2001

Multiscale Computational Tools for Elementary Particles

Achi Brandt
Weizmann Institute of Science

Email: achi.brandt@weizmann.ac.il

Abstract:

Current calculations of quantum fields are extremely expensive and obtain only very poor predictions of elementary particle properties. The enormous cost is the product of several separate complexity factors, which will be explained. In a series of recent studies of simple models, it has been shown that each of these complexity factors can be eliminated or drastically reduced by several new types of multiscale algorithms, including those described below.

A new algebraic multigrid (AMG) technique can solve Dirac Equations on highly disordered gauge fields at linear complexity. The "bootstrap AMG," an approach for fast-solving general discretized PDE systems, will be described in detail.

A special multiscale structure for describing the main part of the inverse of a differential operator allows very fast self-updating upon local changes in the operator. This permits instant updating of the inverse and the operator dominant during Monte Carlo simulations of the gauge field, yielding fast calculations of the fermionic interactions.

A new Monte Carlo approach that combines renormalization and multigrid ideas has been shown to eliminate both the critical slowing down and volume factors, implying that thermodynamic limits can generally be calculated to accuracy *e* in O(e**(-2)) computer operations. Related renormalization-multigrid (RMG) procedures are being developed for fast simulations in statistical mechanics and molecular dynamics.

Speaker's web page: http://www.wisdom.weizmann.ac.il/~achi/

Institution web page: http://www.weizmann.ac.il/

June 12, 2001

ACTC Tools for Application Performance Analysis of Scientific Programs

Luiz De Rose
IBM TJ Watson Research Center

Email: laderose@us.ibm.com

Abstract:

In this talk I will present an overview of the application performance analysis tools under development at the Advanced Computing Technology Center for tuning and optimization of applications running on the IBM platforms. These tools were designed to help users understand the behavior of applications on complex parallel environments. They are based on dynamic instrumentation, access to hardware performance counters, hints to allow users to correlate the behavior of the application to hardware components, and careful mapping of performance data to source code constructs. The presentation will focus on the Hardware Performance Monitor (HPM) Toolkit and the Simulator Guided Memory Analyzer (SiGMA).

The HPM Toolkit consists of libraries and utilities to support performance data capture, analysis, and presentation of hardware performance metrics from applications written in Fortran, C, and C++, executing on sequential or parallel systems. These metrics allow users to correlate the behavior of the application to one or more of the components of the hardware, providing hints that help users to identify causes of performance problems.

SiGMA is a data-centric tool under development that will be able to predict performance and identify bottlenecks, problems, and inefficiencies in a program, due to data layout in the memory hierarchy, as well as propose solutions to improve program performance in current and future architectures.

Institution web page: www.research.ibm.com/actc/

June 8, 2001

The Aqua Approximate Query Answering System

Phillip Gibbons
Bell Laboratories

Email:
gibbons@research.bell-labs.com

Abstract:

Exploratory data analysis (also called OLAP, decision support, and data mining) can be quite time consuming for large data recording and warehousing environments (e.g., sales transaction databases, call detail repositories, customer service historical data). Responses to queries can take minutes to hours, even on top-of-the-line database systems, due to the large amount of computation and disk I/O required to compute an exact answer over a large database.

Aqua is the first system to provide fast, highly accurate approximate answers for a broad class of queries arising in data warehousing environments. Aqua answers queries using precomputed synopses of the data. Approximate answers with error guarantees (e.g., 384K +/- 1K) are provided by rewriting the user query over the synopses and executing the new query. Aqua provides answers in orders of magnitude less time than it takes to compute an exact answer. To avoid costly overheads, Aqua synopses are incrementally maintained in the presence of database updates. Besides the overall system architecture, the key to Aqua is the design and analysis of its synopses. We give examples of novel synopses developed for Aqua; these are also applicable to other "data streams" contexts such as network monitoring.

Aqua is a middleware software solution supporting standard interfaces based on ODBC, to work with any report generating software in the front-end and any commercial DBMS in the back-end. It is currently available for the Windows and Solaris (UNIX) platforms.

Speaker's web page: http://www1.bell-labs.com/user/pbgibbons/

Research web page: http://www.bell-labs.com/project/aqua/

Institution web page: http://www.bell-labs.com/

June 6, 2001

Memory Profiling on Shared-Memory Multiprocessors

Jeff Gibson
Stanford University

Email: jeffg@powderkeg.stanford.edu

Abstract:

Tuning application memory performance can be difficult on any system, but it is particularly so on shared-memory multiprocessors due to the implicit nature of communication and the unforeseen interactions among the processors. Tools, called memory profilers, that allow the user to map memory behavior back to application data structures can be invaluable aids to the programmer. Unfortunately, memory profiling is difficult to implement efficiently since most systems lack the requisite hardware support.

This work describes several techniques that can be used for efficient memory profiling. Each requires differing levels of hardware support, either on the compute processor or on the system node controller. We prototype each system on the versatile FLASH multiprocessor and present measurements of both overhead and accuracy for each method.

Speaker's web page: http://www-flash.stanford.edu/~jeffg/

Research web page: http://www-flash.stanford.edu/

Institution web page: http://www.stanford.edu/

June 4, 2001

Efficient Shared Memory Support in NUMA-based Cluster Environments

Martin Schulz
Technical University, Munich

Email: schulzm@in.tum.de

Abstract:

NUMA-based clusters provide an efficient and cost-effective platform for parallel computing. However, a comprehensive software infrastructure with support for all major parallel programming paradigms is required in order to fully exploit their potentials. Such an infrastructure has been developed within the SMiLE project (http://smile.in.tum.de/) for commodity clusters based on the NUMA interconnection fabric SCI (Scalable Coherent Interface).

While support for message passing can be achieved on top of these architectures in a straightforward manner, the implementation of a Shared Memory programming framework imposes additional challenges. Most severely are the semantic gap between the global physical memory provided by the NUMA hardware and the global virtual memory required for Shared Memory programming and the need for an explicit cache and consistency management.

The HAMSTER (Hybridism-based Adaptive and Modular Shared memory archeTEctuRe, http://hamster.in.tum.de/) framework aims to overcome these challenges and thereby contribute to a multi-paradigm software infrastructure efficiently supporting both message passing and shared memory. Its core component is a Hybrid-DSM system combining the NUMA hardware capabilities with a lean software component for memory and consistency control. A prototype has been implemented within SMiLE and has been evaluated using several applications, including two real-world codes from the area of nuclear medical imaging. The results of these experiments are very encouraging and prove both the feasibility and the efficiency of the proposed approach.

Speaker's web page: http://wwwbode.in.tum.de/~schulzm/

Institution web page: http://wwwbode.in.tum.de/

June 4, 2001

ARIMA Time Series Dynamic Modeling and Forecasting for Adaptive I/O Prefetching

Nancy Tran

University of Illinois, Urbana-Champaign

Email: n-tran1@cs.uiuc.edu

Abstract:

Variability in application data access patterns, together with fluctuations in processor loads and network delays, make input/output (I/O) request arrivals irregular and bursty. Such bursty patterns, combined with transfer-limited storage devices, create a major I/O bottleneck. One approach at relieving this bottleneck is to anticipate future I/O bursts, adaptively prefetching data before it is needed. To be effective, prefetches must be initiated early enough to avoid I/O stalls, but not too early, to minimize unnecessary cache buffer replacement. In this talk, we will introduce the use of time series analysis to dynamically model and forecast diverse I/O temporal patterns during application execution. The forecasts are then used to prefetch read requests during computation intervals to hide I/O latency.

We will describe the design of a "just-in-time" prefetcher. This prefetcher combines time series forecasts with the Pablo group's Markov model spatial forecasts to adaptively determine when, what, and how much to prefetch. Experimental results with I/O intensive scientific codes show improvement in total execution time and cache misses over the traditional Unix file system.

Research web page: http://www-pablo.cs.uiuc.edu/

Institution web page: http://www.uiuc.edu/

May 31, 2001

Immediate Access to Highly Complex, Remote, 4D Models

Jack Snoeyink and Jarek Rossignac

University of North Carolina, Chapel Hill, and Georgia Institute of Technology

Abstract:

Part I: Compression and progressive transmisison of meshes of triangles and tetrahedra (J. Rossignac)

After developing IBM's "Topological Surgery" technique with G. Taubin, which is now the core of the MPEG-4 standard for 3D compression, Jarek Rossignac has focused his research efforts on a simpler compression scheme, called "Edgebreaker." Edgebreaker holds the record for the lowest guaranteed upper bound for encoding planar triangle graphs (1.8 bits per triangle). In practice, it compresses the connectivity of triangle meshes down to about a bit per triangle. A new implementation of Edgebreaker, developed in collaboration with A. Szymczak, A. Safonova, H. Lopes, and V. Coors, works as a simple state machine. The detailed code fits on three pages and uses, as sole data structure, two arrays of integers, called the Corner Table.

The current implementation, which is publicly available, uses a parallelogram prediction but can be easily combined with other schemes for compressing vertex location. Both the Topological Surgery and Edgebreaker have received Best Paper awards. Rossignac and his colleagues have also developed a similar approach, called Grow&Fold, for compressing tegrahedral meshes, such as those used for finite element analysis. Rossignac has also developed progressive techniques for transmitting compressed versions of both triangle meshes and of tegrahedral meshes (with R. Pajarola and A. Szymczak), such that the client may download low resolution models first and then, when necessary, refine them to higher accuracy by downloading compressed upgrades.

Part II: Isosurfaces and contour trees (J. Snoeyink)

Scientific simulations produce data in the form of sample points with intensity values, which are often visualized by drawing level sets or isosurfaces (surfaces of points with the same intensity). One tool that can help in choosing threshold values for interactive exploration of such data is the contour tree, which has been developed by a number of people including van Kreveld, Bajaj, and Pascucci. Contour trees encode the evolution of isosurfaces as the threshold parameter varies. It can be used to compute seed sets for tracing whole or partial isosurfaces, to determine important values for the threshold parameter, and to support flexible contouring. Carr, Snoeyink, and Axen developed a simple algorithm to compute a contour tree without computing the set of all isosurfaces, which enables the use of the contour tree as a data analysis tool.

Part III: Towards a real-time remote visualization of time-dependent 3D simulations (J. Rossignac)

Engineering and scientific simulations conducted at high-performance computing centers produce hundreds of gigabytes of data, which represent the evolution of several variables over a 4D space-time domain. New compression and progressive transmission techniques are needed to support the interactive

(continued next page)

Abstract (continued):

Jack Snoeyink and Jarek Rossignac

University of North Carolina, Chapel Hill, and Georgia Institute of Technology

exploitation of this data. In collaboration with J. Snoeyink and his colleagues at UNC and with A. Szymczak and S. Menon (Aerospace) at Georgia Tech, Jarek Rossignac has been exploring new approaches for the compression, progressive transmission, and interactive visualization of such data sets. The techniques under consideration view the data as a hyper-terrains in five dimensions.

The natural sampling of the hyper-terrain on a regular 4D space-time grid may be concisely encoded at several levels of resolution using higher-dimensional wavelets or other predictive schemes coupled with variable-length coding. However, in interactive applications where selected subsets of the data must be accessed at specific resolutions, regular representations may prove less effective than pentatope meshes produced through an adaptive simplification process or than semi-regular multi-resolution representations. We plan to evaluate all three approaches in the context of a client/server system, where the operator uses two parameters (say pressure P and time T) to interactively control a color-coded iso-surface S(T,P) of all 3D points that take pressure P at time T.

The relevance of each iso-surface, measured for example in terms of its surface area, curvature integral, or number of connected components, may be displayed as the height or color of a P-T control terrain, upon which the (P,T) locations of previously inspected iso-surfaces may be traced and annotated. The global view of the entire dataset provided by the control terrain may help focus interactive inspection on the relevant subsets of the data. We plan to extend to these 4D models the compression and progressive transmission techniques that we have developed for 3D meshes, so that a crude approximation of an initial isosurface S(P,T) may be quickly downloaded and then refined in real time either to follow user-controlled changes in P or T, or to increase its accuracy during navigation pauses.

Part IV: An interface prototype (J. Snoeyink)

We have implemented an initial prototype for the user interaction on the client with Lutz Kettner, a postdoctoral researcher at UNC Chapel Hill. We illustrate this prototype on data sets from combustion simulation, fluid flow, pollution studies, and heat convection in the earth. The data dimensions on the client side are partitioned into a 3D viewer for isosurfaces and a 2D control plane where each point selects a particular isosurface in 3D. Annotations in the control plane help the user to navigate the volume data. We use the number of connected components of the isosurface as an example of an automatically generated annotation that we can compute efficiently using contour trees. We also provide a small pre-computed preview window that is shown continuously at the current cursor position in the control plane. The small preview facilitates a fast overview; selecting the current location in the control plane creates the corresponding detailed view in the 3D viewer for isosurfaces.

May 22, 2001

Higher-Order Semi-Implicit Projection Methods for the Incompressible Navier–Stokes Equations

Michael Minion
University of North Carolina

Email: minion@amath.unc.edu

Abstract:

This talk presents a detailed discussion of the appropriate boundary conditions for semi-implicit projection methods for the incompressible Navier–Stokes equations. In particular, the boundary conditions for the intermediate velocity (the solution of the implicit treatment of the momentum equation) will be analyzed. The relationship between the form of the momentum equation, the boundary conditions for the intermediate velocity, and the pressure equation for second-order projection methods will be made clear. Also, a discussion of higher-order methods based on Spectral Deferred Corrections will be presented.

Institution web page: http://www.northcarolina.edu/

May 18, 2001

Parallel Communication & MPI-I/O Benchmarking and Automatic MPI Profiling

Rolf Rabenseifner
University of Stuttgart

Email: rabenseifner@hlrs.de

Abstract:

The talk will start with an overview on German HPC centers and on the High Performance Computing Center Stuttgart (HLRS) that has systems totaling more than 700 GFlops (Rmax), including new vector machines from NEC and Hitachi.

The second part of the talk describes our automatic counter profiling of all MPI applications on our Cray T3E-900/512. A detailed summary of the hardware performance counters and the MPI calls of any MPI production program is gathered during execution and written on a special syslog file. User specific summaries are sent by e-mail to each user each week. Summary results of profiling our users for six months are presented.

The third part of the talk describes two MPI benchmarks created to characterize the balanced system performance of high-performance clusters and supercomputers: b_eff, the communication specific benchmark examines the parallel message passing performance of a system, and b_eff_io, which characterizes the effective I/O bandwidth. Both benchmarks have two goals: (a) to get a detailed insight into the performance of different parallel communication and I/O patterns, and (b) to obtain a single average bandwidth number. Results are given for IBM SPs, Cray T3E, NEC SX-5, and Hitachi SR 8000.

Speaker's web page: http://www.hlrs.de/people/rabenseifner/

May 16, 2001

Cartesian Grid Embedded Boundary Finite Difference Methods for Partial Differential Equations in Irregular Geometries

Philip Colella

Lawrence Berkeley National Laboratory

Email: colella@colella2.lbl.gov

Abstract:

In this talk, we give an overview of a set of methods being developed for solving classical PDEs in irregular geometries, or in the presence of free boundaries. In this approach, the irregular geometry is represented on a rectangular grid by specifying the intersection of each grid cell with the region on one or the other side of the boundary. This leads to a natural conservative discretization of the partial differential operator on either side of the boundary, with the solution itself being defined on a rectangular mesh. This representation of finite differences in irregular geometries is very appealing for a variety of reasons. The problem of grid generation has already been solved. More generally, the regular geometric structure of rectangular grids simplifies a variety of issues, including control of the truncation error, the development of efficient iterative methods based on multigrid, the introduction of adaptive meshes, and the coupling to other physical submodels, such as particles or radiation. In order to understand these methods, we want to place them in a more systematic numerical analysis setting. Among the tools we will use are truncation error analysis using modified equation models with singular right-hand sides, and the development of specialized discretizations that maintain uniform stability and/or conditioning in the presence of arbitrarily small control volumes.

Speaker's web page: http://seesar.lbl.gov/anag/staff/colella/colella.html

Research web page: http://seesar.lbl.gov/anag/

Institution web page: http://www.lbl.gov/

May 15, 2001

UPS: Simplifying the Performance on Clusters of SMPs

Richard Barrett
Los Alamos National Laboratory

Email: rbarrett@lanl.gov

Abstract:

Large-scale physics simulations must be designed so that they execute effectively in a distributed memory, parallel processing computing environment. Attempts have been made to hide this environment from the application writer (for example High Performance Fortran (HPF)). Unfortunately, performance has been disappointing for many complex programs. Thus far, application writers seeking high performance have been forced to use explicit message-passing software, which has sometimes been referred to as the assembly language of parallel computers. UPS is designed to provide an approach in between these two extremes.

UPS requires that the programmer confront the computing environment, while abstracting away the complexities necessarily associated with it. The result is that the code physicist can use a simplified coding style, natural to the application, which minimizes the time spent moving the data among the distributed processes. The UPS programming model focuses on applications that can be posed using collective operations. In the basic sense this includes broadcasts and reductions of distributed data and barriers. From a higher perspective, this includes gather/scatter capabilities, which is especially useful for unstructured grid applications. Linear solver capabilities (Krylov and AMG methods) have also recently been added.

In this talk I will describe the functionality provided by the UPS library, will discuss our plans for the future, specifically within the context of anticipated computing platforms, and will attempt to determine if collaboration between us would be fruitful.

Speaker's web page:

http://www-xdiv.lanl.gov/XCI/PROJECTS/UPS/rbarrett/rbarrett.html

Research web page: http://www-xdiv.lanl.gov/XCI/PROJECTS/UPS/

Institution web page: http://www.lanl.gov/

May 11, 2001

Approximate Query Answering Using Wavelets

Rajeev Rastogi
Bell Laboratories

Email:
rastogi@research.bell-labs.com

Abstract:

Approximate query processing has recently emerged as a viable, cost-effective solution for dealing with the huge data volumes and stringent response time requirements of today's Decision Support Systems (DSS). Most work in this area, however, has so far been limited in its query processing scope, typically focusing on specific forms of aggregate queries. Furthermore, conventional approaches based on sampling or histograms appear to be inherently limited when it comes to approximating the results of complex queries over high-dimensional DSS data sets.

In this talk, we propose the use of multi-dimensional wavelets as an effective tool for general-purpose approximate query processing in modern, high-dimensional applications. Our approach is based on building wavelet-coefficient synopses of the data and using these synopses to provide approximate answers to queries. We develop novel query processing algorithms that operate directly on the wavelet-coefficient synopses of relational data, allowing us to process arbitrarily complex queries entirely in the wavelet coefficient domain. This, of course, guarantees extremely fast response times since our approximate query execution engine can do the bulk of its processing over compact sets of wavelet coefficients, essentially postponing the expansion into relational tuples until the end-result of the query.

Speaker's web page: http://www.bell-labs.com/user/rastogi/

Research web page: http://www.bell-labs.com/project/serendip/

Institution web page: http://www.research.bell-labs.com/

May 9, 2001

Parallel Adaptive Multigrid Methods with Application to LES

Stefan Lang and Sandra Nagele
University of Heidelberg

Abstract:

Advanced parallel applications based on the message passing paradigm are difficult to design and implement, especially when solution adaptive techniques are used and three-dimensional problems on complex geometries are faced. Many of these difficulties are addressed inside the UG software platform, e.g., dynamic load migration and load balancing, parallel grid adaption, basic grid management, and parallel IO. The main idea during the design process of UG, a multigrid code for the computation of partial differential equations, was to find proper abstractions for each of the different functionality parts of a parallel, adaptive, and unstructured software system. This ensures a maximal degree of code reuse. Therefore, the treatment of various partial differential equations is possible without superfluous coding. In a parallel context grid, adaption involves the need to rebalance the computational load. This stage involves both determining a new load balancing and dynamically redistributing the objects of the grid parts. A key feature of UG is the capability to dynamically migrate grid objects between the processors during run time. This difficult task is supported by DDD (dynamic distributed data), a new parallel programming model. Parallel adaptive calculations of complex real-world problems, flow in porous media and elastoplasticity, based on multigrid methods are presented.

The use of multigrid methods in conjunction with a Large Eddy Simulation (LES) is very compelling since both are based on multiple scales. LES is a turbulence model that resolves large turbulent scales and models the small ones. The scale separation is performed by applying a spatial convolution operator (filter operator) to the incompressible Navier–Stokes equations. As the filter operator, a top hat filter is applied with a grid-dependent support size. Hence, the application of the filter results in a locally varying average in space. The LES model itself also depends on the grid size since the filtering process removes all subgrid scales. Some special dynamic LES models have been developed by various researchers, which apply two filters with different support size at each point of the domain to the governing equation system. Through comparison of the two different large-scale resolutions, the model term can be specified locally. This is similar to the multigrid cycle where the defect is restricted to the coarser grid and higher frequencies are removed. Another property of dynamic models is their ability to adjust themselves to local flow structures. This adaptivity is very useful since in some regions of the domain the flow can be laminar and a turbulence model is not necessary at all. Hence dynamic models were used in the simulations.

Since unstructured grids are used, the filter width varies in the solution domain. Using unstructured grids the resolution of the turbulent scales can be increased locally to decrease the modeling effort. In the neighborhood of walls, for example, it is possible to use a smaller grid size. By this grid adaptation the local flow structures can be better resolved and less modeling with less modeling error is necessary.

A Krylov subspace method with linear multigrid as preconditioner is used to solve the linearized system. Within the multigrid cycle it is important to separate modeling and solving in the sense that only on the finest grid the modeled part of the equations is determined as described above. This is necessary for an appropriate modeling of the subgrid turbulent scales. Afterwards the model term is restricted to the coarse grids and a standard linear multigrid cycle can be used. This solution strategy was applied to different flow problems, which will be presented.

Institution web page: http://www.uni-heidelberg.de/

May 8, 2001

Low Mach Number Compressible Flow Simulations

Achim Gordner
University of Heidelberg

Abstract:

Flow simulations that take into account variable density are of importance in modeling laminar flow combustion. In recent laminar flame simulations, we used an approach based on rigorous low Mach number asymptotics. In the zero Mach number limit the acoustic modes are eliminated. However, for some applications it is necessary to take the effects of acoustic waves into account. Since low Mach number computations are considered, a discretization scheme for incompressible flow equations has been extended to the isotropic compressible case. Hence, the approach is using primitive variables for the discretization of the Euler equations. For very low Mach numbers the resulting algebraic system of equations becomes stiff, which makes further demands on the smoother used in the multigrid cycle. Preliminary results will be shown.

Institution web page: http://www.uni-heidelberg.de/

May 8, 2001

Numerical Simulation in Science and Engineering

Gabriel Wittum
University of Heidelberg

Abstract:

Numerical simulation has become one of the major topics in computational science. To promote modeling and simulation of complex problems, new strategies are needed allowing for the solution of large, complex model systems. Crucial issues for such strategies are reliability, efficiency, robustness, usability, and versatility.

After discussing the needs of large-scale simulation we point out basic simulation strategies such as adaptivity, parallelism, and multigrid solvers. These strategies are combined in our simulation system UG (for “Unstructured Grids”).

In the second part of the seminar we show the application of these strategies to the simulation of processes from fluid dynamics and subsurface modeling. In particular we will discuss the construction of the Filtering Algebraic Multigrid methods (FAMG) and show an application of it for the simulation of bioremediation of a chlorine-contaminated aquifer.

Institution web page: http://www.uni-heidelberg.de/

May 4, 2001

Recent Advances in Direct Numerical Simulations (DNS) of Turbulent Shear Flows Laden with Particles (droplets or bubbles)

Said Elghobashi
University of California, Irvine

Email: selghoba@uci.edu

Abstract:

Particle-laden turbulent flows are encountered in nature and numerous engineering applications. Sand and dust storms near the surfaces of Earth and Mars particulate pollutants in the atmosphere and droplet sprays in combustors are some examples. Recent advances in parallel computer technology allowed the direct numerical solution (DNS) of the governing equations of these flows with minimum approximation. There are at present two formulations of DNS of particle-laden flows: Eulerian-Lagrangian (or trajectory) and Eulerian-Eulerian (or two-fluid). When the volume fraction of the dispersed particles is of the order of 0.0001 or greater, the particles modify the turbulence structure; this phenomenon is known as two-way coupling.

The objective of this seminar is to present an overview of the above two approaches and DNS results for some particle-laden and bubble-laden turbulent flows.

Speaker's web page: http://kolmog.eng.uci.edu/

Research web page: http://mae.eng.uci.edu/

Institution web page: http://www.uci.edu/

April 27, 2001

Global Data Distribution in Software Distributed Shared Memory Systems

David Lowenthal
University of Georgia

Email: dkl@cs.uga.edu

Abstract:

Distributing data is one of the key problems in implementing efficient distributed-memory parallel programs. The problem becomes more difficult in programs where data redistribution between computational phases is considered. The global data distribution problem is to find the optimal distribution in multi-phase parallel programs. We are investigating this problem in the context of a software distributed shared memory (SDSM) system.

In this talk, we first describe SUIF-Adapt, which is an integrated compiler/run-time system for finding efficient global data distributions, and then we discuss performance of SUIF-Adapt programs. We next describe how SUIF-Adapt obtains accurate redistribution time estimates. Obtaining these estimates is difficult because it depends on access patterns, page locations, and the SDSM consistency protocol. However, having accurate redistribution times is critical to choosing an efficient distribution. We discuss our integrated compiler/run-time method for finding accurate time estimates and describe optimizations that improve the performance of our algorithm.

Speaker's web page: http://www.cs.uga.edu/~dkl/

Institution web page: http://www.uga.edu/

April 17, 2001

A Survey of Some Recent Results on Two-Level Finite Element Preconditioning Methods

Owe Axelsson
University of Nijmegen

Email: axelsson@sci.kun.nl

Abstract:

Two-level methods for conforming finite element approximations of self-adjoint scalar or systems of elliptic equations are considered. Although this topic was first treated some twenty years ago, there are some recent results and there remain some open questions. The matrix that arises is naturally partitioned in a two-by-two block form $A=[Aij]$, $i,j=1,2$. The simplest but still powerful preconditioner is the block diagonal part, i.e., B = blockdiag (A11, A22). The corresponding spectral condition number is $(1 + \tilde{a})/(1 - \tilde{a})$, where $\tilde{a}$, $0 < \tilde{a} < 1$ is the constant in the strengthened CBS (Cauchy–Buryakowski–Schwarz) inequality. Equivalently, $\tilde{a}$ is the cosine of the angle between the subspaces spanned by the coarse mesh basis functions and the basis functions for the added node points, respectively.

The method can be extended by recursion to a multilevel method by approximating the coarse mesh matrix (A22) in a similar way. In this way, the resulting condition number can be stabilized using certain matrix polynomial approximations of Schur complement matrices.

Two major partly open questions arise: (1) the computation of the constant $\tilde{a}$ for three-dimensional bodies and the dependence of this constant on the diffusion coefficients; (2) preconditioning of the block matrix corresponding to the added degrees of freedom on each level to achieve a condition number which is bounded uniformly in the problem parameter and the level number.

Examples of some new recent results on each topic will be presented and some open questions will be discussed.

Speaker's web page:

http://www-math.sci.kun.nl/math/medewerkers/Axelsson.html

Institution web page: http://www.kun.nl/

April 9, 2001

A Review of Computational Modeling in the MHD Laboratory at STRELA

Oleg Diyankov
STRELA Open Computer Center

Abstract:

We present three topics in computational physics. The first is a brief description of the MHD model used for "2.5D" modeling. The mathematical approach to the solution of the 2.5D MHD equations will be discussed. Application of our code to Z-Pinch implosion problems and the treatment of magnetic field influence on laser-produced plasma motion will be presented. The second part of the talk is devoted to the description of a 2D free Lagrangian code. The mathematical approach to the solution of the Lagrangian gas dynamic equations on a moving grid, consisting of Voronoi cells, will be discussed. Example problems solved with the use of the code will be presented. The third part concerns the models and ideas that we would like to implement in the frame of an ISTC project, "Computer Modeling of Ceramic Pressing—the Creation of Ultra-high-hardness Materials." Our ideas on high order difference schemes for irregular grids will be discussed.

April 9, 2001

Numerical Simulations of Flows with Two Viscous Liquids

Yuriko Renardy

Virginia Polytechnic Institute and State University

Email: renardyy@math.vt.edu

Abstract:

We use a volume-of-fluid method to track the deformation and breakup of the interface separating two viscous liquids. The original code is Stephane Zaleski's SURFER. To this we have added a semi-implicit Stokes solver to handle low Reynolds numbers, parallelization to handle large 3D simulations, and new physics such as a model for a yield stress liquid, a model for surfactants, and moving contact lines.

Two simulations will be described. One is the axisymmetric simulation of bamboo waves for core-annular flow of oil and water. The second is rupturing of a viscous drop under simple shear including the effect of inertia. Videos will be shown.

Speaker's web page: http://www.math.vt.edu/people/renardyy/

Institution web page: http://www.vt.edu/

April 9, 2001

Textbook Multigrid Efficiency in Solution of Reynolds-Averaged Navier–Stokes Equations

Boris Diskin
ICASE, NASA Langley

Email: bdiskin@icase.edu

Abstract:

Full multigrid (FMG) algorithms are the fastest solvers for elliptic problems. These algorithms can solve a general discretized elliptic problem to the discretization accuracy in a computational work that is a small (less than 10) multiple of the operation count in one target-grid residual evaluation. Such efficiency is known as textbook multigrid efficiency (TME). The difficulties associated with extending TME for solution of the Reynolds-averaged Navier–Stokes (RANS) equations relate to the fact that the RANS equations comprise a system of coupled nonlinear equations that is not, even for subsonic Mach numbers, fully elliptic, but contains hyperbolic parts. TME for the RANS simulations can be achieved if the different factors contributing to the system can be separated and treated optimally, e.g., by multigrid for elliptic factors and by downstream marching for hyperbolic factors. One means of separating the factors is the distributed relaxation approach proposed by A. Brandt. Earlier demonstrations of TME solvers with distributed relaxation have been performed for incompressible free-stream inviscid and viscous flows without boundary layers.

In this talk, a general framework for achieving TME in solution of the high-Reynolds-number Navier–Stokes equations will be outlined. TME distributed-relaxation solvers will be demonstrated for viscous incompressible and subsonic compressible flows with boundary layers and for inviscid compressible transonic flows.

Speaker's web page: http://www.icase.edu/~bdiskin/

Institution web page: http://www.icase.edu/

March 20, 2001

Boundary Conditions and Estimates for the Stokes Equations on Staggered Grids

Bertil Gustafsson
Stanford University

Email: bertil@sccm.stanford.edu

Abstract:

We consider the steady state Stokes equations, describing low speed flow, and derive estimates of the solution for various types of boundary conditions. When using the equations in their original form with the divergence condition as part of the system, there is a necessary restriction on the boundary data. We formulate the boundary conditions in a new way, such that this problem is eliminated, and the boundary value problem becomes nonsingular and well conditioned. By using a difference approximation on a staggered grid, we are able to derive a nonsingular approximation in a direct way, and such that we get the same type of estimates as for the continuous case. Numerical experiments confirm the theoretical results.

In a recent thesis by Jonas Nilsson, the method is generalized to the time-dependent problem, to curvilinear grids, and to the incompressible Navier–Stokes equations. We will present a brief survey of this work as well.

Institution web page: http://www.stanford.edu/

March 14, 2001

Fast High-Order Methods in CEM

Oscar Bruno
California Institute of Technology

Email: bruno@acm.caltech.edu

Abstract:

I will discuss a class of fast, high-order solvers for problems of electromagnetic and acoustic scattering. Our methods include approaches based on perturbation of boundaries and frequencies, as well as novel fast, high-order integral techniques. I will present numerical results for a variety of computational problems containing high frequencies, complex scatterers, and geometric singularities such as edges and corners. (Collaborators: M. Caponi, M. Hyde, L. Kunyansky, F. Reitich, A. Sei)

Speaker's web page: http://www.acm.caltech.edu/%7Ebruno/

Research web page: http://www.acm.caltech.edu/

Institution web page: http://www.caltech.edu/

February 28, 2001

Optimal Preconditioned Eigensolvers for Very Large Symmetric Eigenproblems

Andrew Knyazev
University of Colorado, Denver

Email: andrew.knyazev@cudenver.edu

Abstract:

Many applied and engineering problems, e.g., in structural mechanics, design of nuclear reactors, ocean modeling, and quantum chemistry, lead—after simplifications and an appropriate approximation of original partial differential equations—to extremely large and ill-conditioned linear systems with symmetric positive definite matrices of coefficients and similar symmetric eigenvalue problems. The preconditioned conjugate gradient method became the standard solver for such linear systems. Our ultimate goal is to develop an analogous optimal method for eigenproblems. Ideally, we want to be able to compute an eigenvector of interest at the same cost as that of solving a linear system of equations, using the same preconditioner. That would allow, in particular, a simple adaptation for eigenproblems of available domain decomposition-based and multigrid preconditioners for linear systems.

Searching for the optimal eigensolver, we describe the Locally Optimal Block Preconditioned Conjugate Gradient (LOBPCG) method for symmetric eigenvalue problems, based on a local Rayleigh–Ritz optimization of a three-term recurrence. LOBPCG can be viewed as a nonlinear conjugate gradient method of minimization of the Rayleigh quotient, which takes advantage of the optimality of the Rayleigh–Ritz procedure. Numerical results establish that our LOBPCG method is practically as efficient as the best possible algorithm on the whole class of preconditioned eigensolvers. We discuss several competitors, namely, some inner-outer iterative preconditioned eigensolvers. Direct numerical comparisons with one of them, the inexact Jacobi–Davidson method, show that our LOBPCG method is more robust and converges almost two-times faster. Finally, we show numerically that the LOBPCG method is robust with respect to variable preconditioning. A MATLAB code of the LOBPCG method and the Preconditioned Eigensolvers Benchmarking are available at http://www-math.cudenver.edu/~aknyazev/software/CG/.

Speaker's web page: http://www-math.cudenver.edu/~aknyazev/

Institution web page: http://www.cudenver.edu/

February 16, 2001

Combinatorial Morse Theory and Persistence, with Applications to Distinguishing Noise from Feature

John Harer
Duke University

Email: John.Harer@Duke.edu

Abstract:

This talk will describe a combinatorial approach to analyzing data sets in two and three dimensions. Based on a discrete version of Morse Theory, the approach is first to build a combinatorial framework for the data, and then to work with that framework numerically. This approach is useful in many cases in controlling error accumulation. The talk will also discuss the theory of persistence as a mechanism for distinguishing noise from feature in data sets and will outline procedures for algorithm development and data handling that fit this approach.

Speaker's web page: http://www2.math.duke.edu/faculty/harer/

Institution web page: http://www.duke.edu/

February 9, 2001

Accounting for Stability: Accurately Estimating the Error of Numerical Solutions of Differential Equations

Donald Estep
Colorado State University

Email: estep@math.colostate.edu

Abstract:

Accurately estimating the error of numerical solutions of differential equations remains an important scientific problem. Recently, we have made significant progress using a new approach, which at its heart is computational rather than analytical. This approach is based on a variational a posteriori error analysis that takes into account both the local introduction of discretization error and the accumulation of errors. I will explain the ingredients of this theory and illustrate its application using a variety of examples.

Speaker's web page: http://www.math.colostate.edu/~estep/

Institution web page: http://www.colostate.edu/

January 5, 2001

Exact Analysis of the Cache Behavior of Nested Loops

Erin Parker

University of North Carolina, Chapel Hill

Email: parker@cs.unc.edu

Abstract:

We use Presburger arithmetic to exactly model the behavior of loop nests executing in a memory hierarchy. Our formulas can be simplified efficiently to count various kinds of cache misses and to determine the state of the cache at the end of the loop nest. Our model is powerful enough to handle imperfect loop nests and various flavors of nonlinear array layouts based on bit interleaving of array indices. We also indicate how to handle the modest levels of associativity found in current data caches, and how to perform a certain amount of symbolic analysis of cache behavior. The complexity of the formulas relates to the static structure of the loop nest rather than to its dynamic trip count, allowing our model to gain efficiency in counting cache misses by exploiting repetitive patterns of cache behavior. Validation against cache simulation confirms the exactness of our formulation. Our method can serve as the basis for a static performance predictor to guide program and data transformations to improve performance.

Institution web page: http://www.unc.edu/

January 4, 2001

A Compiler-Based Approach to Specializing Software Libraries

Calvin Lin
University of Texas, Austin

Email: lin@cs.utexas.edu

Abstract:

Software libraries are a popular means of reusing code. Unfortunately, the static black-box nature of libraries can thwart compiler optimizations, limit code reuse, and threaten portability. This talk introduces a new compiler-based approach to optimizing both software libraries and the application programs that use them. The key is a simple annotation language that describes semantic information about libraries. The result is a compiler that can analyze and transform library operations in the same way that language primitives can be analyzed and transformed. Experiments with the PLAPACK parallel linear algebra library show that this technique can yield significant performance improvements, even for a library that has been carefully designed to provide good performance. One goal of this talk is to identify potential users of our system and to meet with programmers who use libraries written in C or Java.

Speaker's web page: http://www.cs.utexas.edu/users/lin/

Research web page: http://www.cs.utexas.edu/users/lin/projects.html

Institution web page: http://www.utexas.edu/

November 30, 2000

A Particle Method and Adaptive Treecode for Vortex Sheet Motion in 3D Flow

Robert Krasny
University of Michigan

Email: krasny@math.lsa.umich.edu

Abstract:

A particle method is presented for computing vortex sheet motion in three-dimensional flow. The particles representing the sheet are advected using the Rosenhead–Moore form of regularized Biot–Savart kernel and new particles are inserted during the simulation to maintain resolution as the sheet rolls up. The particle velocities are computed using a treecode based on Taylor approximation in Cartesian coordinates. The necessary coefficients are computed using a recurrence relation, and several adaptive techniques are employed to gain efficiency, including nonuniform rectangular cells, variable order approximation, and a run-time choice between approximation and direct summation. Tests show that the treecode is significantly faster than direct summation for systems having a large number of particles. The algorithm is applied to simulate flows in which a vortex sheet rolls up to form a vortex ring. Two examples are presented: (1) azimuthal waves on a vortex ring, and (2) merger of two vortex rings.

(This is joint work with Keith Lindsay, NCAR.)

Speaker's web page: http://www.math.lsa.umich.edu/~krasny/

Institution web page: http://www.umich.edu/

November 28, 2000

Checking System Rules Using System-Specific, Programmer-Written Compiler Extensions

Dawson Engler
Stanford University

Email: engler@csl.stanford.edu

Abstract:

Systems software such as OS kernels, embedded systems, and libraries must obey many rules for both correctness and performance. Common examples include, "Accesses to variable A must be guarded by lock B," "System calls must check user pointers for validity before using them," and "Disabled interrupts must be re-enabled." Unfortunately, adherence to these rules is largely unchecked.

This talk shows system implementors can use "meta-level compilation" (MC) to write simple, system-specific compiler extensions that automatically check their code for rule violations. By combining domain-specific knowledge with the automatic machinery of compilers, MC brings the benefits of language-level checking and optimizing to the higher, "meta" level of the systems implemented in these languages. This talk presents results of applying MC to four real systems: Linux, OpenBSD, the Xok exokernel, and the FLASH machine's embedded software. MC extensions found over 600 errors in these systems and led to numerous kernel patches. Most extensions were less than a hundred lines of code and written by implementors who had a limited understanding of the systems checked.

Speaker's web page: http://www.stanford.edu/~engler/

Institution web page: http://www.stanford.edu/

November 27, 2000

About Mesh Generation and Adaptive Multigrid Methods

Joachim Schoeberl
Texas A&M University

Email: joachim@math.tamu.edu

Abstract:

Numerical simulation of complicated processes needs the interaction of different scientific disciplines. Here, the value of software design must not be underestimated. Especially, adaptive multilevel algorithms on complex 3D geometries require a careful design as well as a certain amount of manpower.

The first step in the simulation is the description of the geometry. From the geometry database, the mesh generator automatically builds the initial mesh. To gain optimal advantage of adaptive and hierarchical algorithms, the initial mesh shall be as coarse as possible to resolve the geometry. The geometry database and the mesh database, together with the problem description, are fed into the actual simulation program. Based on the approximate solution, a posteriori error indicators mark elements for refinement or coarsening. Hierarchical mesh refinement algorithms have to perform the changes and maintain information needed for multilevel solvers. It is important to have access to the geometry to adapt the mesh to the boundary and edges.

We will present available tools for the generation of the initial mesh and for performing the necessary mesh refinement. The underlying geometry description is based on a Constructive Solid Geometry (CSG) tree and includes now various types of spline curves and surfaces as well. The mesh generator NETGEN performs the steps from scanning the geometry until generation of the volume mesh. Small details in the geometry are detected and the mesh is automatically graded. This mesh can be used by any simulation program, which also sets the refinement flags. NETGEN then performs the necessary updates and returns the mesh together with the hierarchical information to the simulation code. There are different realizations of the interface, from file transfer to direct data access.

We will present applications, where NETGEN is used together with our fem code for the simulation of mechanical, electromagnetical, as well as coupled field problems.

Speaker's web page: http://www.math.tamu.edu/~joachim/

Institution web page: http://www.tamu.edu/

November 16, 2000

Janus: A Library for Adaptive Unstructured Scientific Applications

Peter Gottschling
GMD First

Email: pg@first.gmd.de

Abstract:

Janus is a conceptual framework and a C++ template library for irregular and regular scientific applications. It provides (potentially distributed) data structures to represent spatial structures and numerical data of these applications. Janus is implemented on top of the C++ Standard Template Library and uses MPI as default parallel platform. The components of Janus cooperate well with existing linear solvers and mesh partitioning packages, such as PETSc and ParMetis.

Janus rests on the observation that there occur two essential kinds of objects in scientific applications. The first kind is referred to as "spatial structures" such as (rectangular) grids, meshes, or graphs. The second kind is simulation data that are associated with the spatial structures. Typical examples are grid functions (vectors), element matrices on finite element meshes, and other (sparse) matrices. An important observation is that the spatial structures are conceptually more stable than the associated data. Janus takes advantage of this stability by putting as much computational effort as possible (load balancing, setting local index schemes, etc.) in the construction of spatial structures. On the other hand, the access to the spatial structures and the usage of associated data are fast.

The conceptual framework of Janus is designed using the paradigm of generic programming that has been successfully applied in the C++ standard library and other libraries such as the Matrix Template Library (MTL) or the POOMA framework. Using templates enables a more problem-oriented programming style, i.e., domains of arbitrary types, like triangles or hexahedra, can be defined. Furthermore, relations are defined between arbitrary domains, not between index sets.

For more information about Janus visit: http://www.first.gmd.de/promise/

Research web page: http://www.first.gmd.de/~jens/janus/

Institution web page: http://www.first.gmd.de/Welcome.ORIG.html

November 2, 2000

Virtual Global Grids for Adaptive Methods in Space and Time

Ulrich Ruede
University of Erlangen-Nuernberg

Email: ruede@informatik.uni-erlangen.de

Abstract:

Many advanced methods use adaptive techniques. In this talk we will present a time-space adaptive discretization and iterative solution technique. The method is based on the fully adaptive multigrid method (FAMe) and virtual global grids, which have been proposed for the adaptive solution of elliptic equations. Here, adaptive discretization and iterative solution are integrated in one process. The approach can also be extended to efficiently exploit locality as needed, e.g., for an efficient parallelization or cache-aware algorithms.

In the time-dependent case, we propose not to use global time steps, but to treat time with a multilevel approach, similar to time parallel multigrid methods. Thus, the iterative solver iterates on all time levels simultaneously. We will discuss how this can be extended to use adaptive meshes in space and time. This method has been implemented and tested for prototype model problems.

Speaker's web page: http://www10.informatik.uni-erlangen.de/~ruede/

Institution web page: http://www.uni-erlangen.de/

October 26, 2000

Algebraic Multilevel Preconditioning Based on Approximate Cyclic Reduction

Johannes Kraus
University of Leoben

Email: jkraus@poptop.llnl.gov

Abstract:

We consider preconditioners for large sparse matrices arising from discretization of partial differential equations of predominant elliptic type. The proposed purely algebraic multilevel method is based on approximate cyclic reduction. Within an incomplete LU, decomposition process spanning trees of matrix graphs are constructed that rest on a local optimization principle. A red-black coloring of these subgraphs yields the partitioning into fine- and coarse-grid variables and is also utilized to determine appropriate approximations of the Schur complements (the coarse-grid operators) on different coarse levels. This idea is combined with algebraic multilevel iterations, V- or W-cycle, in the preconditioning step of some Krylov subspace method. Another variant defines a nonlinear algebraic multilevel iteration of W-cycle type based on inner GCG iterations. This results in a parameter-free method that shows similar convergence properties and optimal order of computational complexity for various model problems.

Institution web page: http://www.unileoben.ac.at/

October 24, 2000

Semi-Unstructured Grids in 3D

Christoph Pflaum
University of Wurzburg

Email: pflaum1@llnl.gov

Abstract:

It is well known that structured grids have several advantages in comparison to unstructured grids (e.g., high performance, low storage requirement, local refinement, and construction of coarse grids). The disadvantage of pure structured grids is the difficulty of constructing an automatic grid generator for structured grids on general domains. To avoid this problem one can use semi-unstructured grids. Such grids consist of a maximal structured grid in the interior of the domain and a subdivision of boundary cells by tetrahedra. It will be shown that the number of types of boundary cells in such a construction can be as few as twelve, if the domain does not contain slits. Furthermore, each boundary cell can be subdivided by tetrahedra such that the maximum interior angle is 163 degrees in case of a general domain and 145 degrees in case of a convex domain. Theoretical and numerical results are presented.

October 19, 2000

Osiris Analysis: A Diagnostic and Visualization Package for Large Data Sets from Particle in Cell Codes

Ricardo Fonseca
University of California, Los Angeles

Email: zamb@physics.ucla.edu

E.S. Dodd, R.G. Hemker, S. Lee, C. Ren, F. Tsung, L.O. Silva, W.B. Mori

Abstract:

Visualization is an essential part of any numerical laboratory. The data being produced by present day simulations is both large and complex, and it requires an appropriate set of tools for the adequate understanding of the results obtained. We present here the Osiris Analysis package, developed for the visualization and analysis of scientific data from Particle-In-Cell (PIC) simulations. This package consists of a wide set of visualization routines for scalar and vector data up to three dimensions in size, and it can also be used for any type of grid-deposited data and general particle data diagnostics. The analysis package includes spectral analysis routines, envelope and centroid analysis, *n*-dimensional local peak analysis, and a Poisson solver. This package was developed using the Interactive Data Language (IDL) software package and includes a set of interfacing routines based on the HDF file standard. We will discuss applications from Osiris 3D simulations for the Weibel Instability and Intense Laser/Beam propagation in plasmas.

Speaker's web page: http://exodus.physics.ucla.edu/

Institution web page: http://www.ucla.edu/

October 18, 2000

Algorithm Development for the Stanford University ASCI TFLO Solver

Antony Jameson (with Juan Alonso)
Stanford University

Email: jameson@baboon.stanford.edu

Abstract:

In this talk we describe the efforts, objectives, and methods of the ASCI turbomachinery simulations group at Stanford University. Our group is in charge of the development of efficient solvers for the computation of large-scale unsteady flow simulations in compressors and turbines with multiple stages using a Reynolds Averaged Navier–Stokes model. We begin by presenting a description of our objectives and relating some of the details of the flow solver, TFLO, which is being used for this project. In particular, the dual-time stepping, multigrid-driven algorithm is explained in some detail together with its performance, potential, and shortcomings. We will also present the results of a variety of unsteady large-scale test cases that we are currently conducting on ASCI platforms. Finally, we discuss two alternative approaches that can help speed up the computation of unsteady flows. These approaches are based on a modified ADI/dual-time scheme and on a harmonic balance technique.

Research web page: http://cits.stanford.edu/

Institution web page: http://www.stanford.edu/

October 16, 2000

Diffpack Object-Oriented Software Solutions for PDEs

Are Magnus Bruaset
Numerical Objects AS

Email: amb@nobjects.com

Abstract:

There is rapidly growing interest in the use of modern software technology and object-oriented principles in the context of scientific computing. The C++ software framework Diffpack is the result of more than ten years of dedicated research and development in this direction, with a special emphasis on numerical solution of partial differential equations.

Diffpack is today a commercial product from Numerical Objects in Oslo, Norway. The software is under continuous development in close collaboration with the Scientific Computing group at the University of Oslo.

The talk will focus on the underlying principles of Diffpack and will show several application examples addressing academic as well as industrial problems. These applications are taken from a wide range of areas, including traditional engineering disciplines like structure mechanics, computational fluid dynamics and porous media flow, as well as computational medicine and computational finance.

We will present recent advances in the software, with special emphasis on the Diffpack Parallel Computing Toolboxes. This functionality, which will ship early next year, allows a very easy transition from a sequential to a parallel Diffpack simulator. We will present results from using this software approach in a wide range of applications on a Linux-based cluster of workstations, as well as in traditional parallel environments.

Institution web page: http://www.nobjects.com/

October 5, 2000

Coupled Flows with a Focus on the Eye

Victor Barocas
University of Minnesota, Twin Cities

Email: baroc001@tc.umn.edu

Abstract:

Coupled flows, i.e., those that involve flow of a fluid past a deformable solid, present a number of challenges, but they are extremely common and thus of considerable importance. The greatest challenge associated with the coupled problem is that the domain for the fluid flow problem is not known until one has solved the solid deformation problem, which is in turn coupled to the fluid problem by drag-induced deformation. We use a pseudo-solid approach in which an imaginary compressible solid (the pseudo-solid) is overlaid onto the fluid domain, and the motion of the domain is tracked by monitoring the motion of the pseudo-solid. We have used the pseudo-solid approach to solve the coupled flow problem associated with the passive deformation of the iris (the washer-shaped colored part of the eye) in response to flow of aqueous humor (the clear fluid that bathes the iris). I will present results for the axisymmetric form of the problem with application to healthy and diseased eyes, and I will also discuss our computational time issues (especially for 3D) and our work with T. Manteuffel and S. McCormick of the University of Colorado to develop faster solution methods.

Research web page: http://www.bme.umn.edu/

Institution web page: http://www1.umn.edu/twincities/

ISCR

Institute for Scientific Computing Research

ITS Lecture Series and Lecturer Biosketches

CASC Center for Applied Scientific Computing

July 11, 2001

Computer Simulation of the Heart

Charles S. Peskin

Courant Institute of Mathematical Sciences, New York University

Email: peskin@CIMS.nyu.edu

Abstract:

The study of blood flow in the cardiac chambers is complicated by the interaction of the blood with the flexible heart valve leaflets and with the muscular heart walls. These elastic and contractile tissues have certain properties in common with blood: they are incompressible and neutrally buoyant. These considerations suggest a unified treatment of cardiac mechanics in which the valves and heart walls appear as specialized regions of the fluid in which additional stresses appear. The additional stresses have a particular form, since they are fiber generated, the fibers in question are muscle fibers in the heart walls and collagen fibers in the heart valve leaflets.

We shall describe a computer model of the heart based on these principles and constructed, therefore, as a system of elastic and contractile fibers immersed in viscous incompressible fluid. This virtual heart includes representations of left and right atria; left and right ventricles; mitral, aortic, tricuspid, and pulmonic valves; pulmonary veins and ascending aorta; superior and inferior vena cavae and main pulmonary artery. Sources and sinks in the model veins and arteries connect the heart to pressure reservoirs representing the rest of the circulation. The equations of motion of the virtual heart are solved by the immersed boundary method, and the results are shown as a video animation of the beating heart.

Speaker's web page: http://www.math.nyu.edu/faculty/peskin/

Institution web page: http://www.cims.nyu.edu/

Biographical Sketch:

Charles S. Peskin

Charles S. Peskin's undergraduate studies were in Engineering and Applied Physics (A.B., Harvard, 1968) and his graduate studies were in Physiology (Ph.D., Albert Einstein College of Medicine, 1972). He now combines those interests as a Professor of Mathematics, Courant Institute of Mathematical Sciences, and Member of the Center for Neural Science, at New York University. His teaching at NYU ranges from graduate courses like "Mathematical Aspects of Heart Physiology" to a freshman seminar on "Computer Simulation."

Peskin has worked on several problems in which mathematics and computing are applied to biology and medicine. Some examples are blood flow in the heart, computer-assisted design of prosthetic cardiac valves, fiber architecture of the heart and its valves, fluid dynamics of the inner ear, photon noise in vision and nuclear medicine, and Brownian ratchet dynamics of biomolecular motors. He is the inventor of the immersed boundary method, which is broadly useful for problems of biological fluid dynamics. Peskin is a former MacArthur Fellow. He is a Member of the National Academy of Sciences and also of the Institute of Medicine.

May 9, 2001

Information Security: The Road Ahead

Eugene H. Spafford
Purdue University

Email: spaf@cerias.purdue.edu

Abstract:

Melissa. ILOVEYOU. Web page defacement. Denial-of-service against e-commerce sites. Theft of laptops from the State Department. Information Warfare. Y2K. Napster. Hacking into Microsoft's network. Theft of information at the World Economic Forum. Fraudulent certificates from Verisign.

It is no longer possible to avoid stories of information loss, fraud, and compromise. Open any paper or magazine, or listen to the news (on-line, as well as in the standard media) and stories relating to information security are sure to be present.

So, what is the current state of information security? Are things getting better, or are they getting worse? And what are the challenges that we are likely to see in the near future?

This talk presents some highlights of what is happening in information security, and what is yet to happen. We also include some discussion of the nature of infosec-related challenges that we are likely to face—and few of them are based solely in technology.

Biographical Sketch:

Eugene H. Spafford

Eugene H. Spafford is a professor of Computer Sciences at Purdue University, a professor of Philosophy, and is Director of the Center for Education Research Information Assurance and Security. CERIAS is a campus-wide multi-disciplinary Center, with a broadly focused mission to explore issues related to protecting information and information resources. He has written extensively about information security, software engineering, and professional ethics. He has published over 100 articles and reports, has written or contributed to over a dozen books, and he serves on the editorial boards of most major infosec-related journals.

Dr. Spafford is a Fellow of the ACM, a Fellow of the AAAS, a Fellow of the IEEE, and a charter recipient of the Computer Society's Golden Core award. In 2000, he was named as a CISSP, honoris causa. Among his many activities, he is chair of the ACM's U.S. Public Policy Committee, a member of the Board of Directors of the Computing Research Association, and a member of the U.S. Air Force Scientific Advisory Board. He was the year 2000 recipient of the NIST/NCSC National Computer Systems Security Award, generally regarded as the field's most significant honor in information security research, and was named as one of the "Five Most Influential Leaders in Information Security" by the readers and editors of Information Security in 1999. In 2001, he was named as one of the recipients of the Charles B. Murphy award, Purdue University's highest award for outstanding undergraduate teaching. In his spare time, Spaf wonders why he has no spare time.

More information may be found at
http://www.cerias.purdue.edu/homes/spaf/

March 28, 2001

Computing the Cosmic Web and the Evolution of the Universe

Michael L. Norman
University of California, San Diego

Email:
mnorman@mamacass.ucsd.edu

Abstract:

Our physical universe is an initial value problem amenable to direct numerical simulation. With known laws of physics and initial conditions, multiresolution numerical algorithms, and sufficient computing power, we are able to simulate the formation and evolution of cosmic structures (galaxies, superclusters, etc.) from shortly after the Big Bang to the present day. In this talk, I will describe how this is done and what we have learned.

Recent astronomical observations have strongly constrained the numerical values of the cosmological parameters governing the global geometry and expansion of the universe, the mean density of its matter constituents, and have strongly constrained the power spectrum of primordial matter fluctuations that seeded the formation of galaxies and clusters. We thus have arrived at a standard cosmological model whose consequences for structure formation are being explored using multiphysics, multiscale numerical simulations on terascale platforms.

A generic feature of these models is that gravity organizes matter into a filamentary network termed the cosmic web. The cosmic web evolves with time nearly self-similarly, with small length and mass scales going nonlinear at early times, followed by larger scales at later times. Structures in the universe thus are predicted to be built up hierarchically "bottom-up," in agreement with observations. In this talk, I will explore the implications of the standard model to the first generation of cosmic structures and their influence on subsequent generations.

The principal computational challenges are twofold: the vast range of length scales that must be resolved simultaneously in a single simulation (multiscale challenge), and the large number of physical processes that must be included for accurate predictions (multiphysics challenge). I describe a decade-long code development quest that has resulted in the world's first adaptive mesh refinement (AMR) application for cosmological structure formation. Optimizing this code for current terascale systems has proved to be an interesting challenge, as I will relate. Visualizing the results of multiscale AMR simulations has required the development of new tools, which we have made publicly available.

Speaker's web page: http://www.ucsd.edu/

Web links:

Laboratory for Computational Astrophysics (http://lca.ncsa.uiuc.edu/lca.html)

LCA Vision (http://zeus.ncsa.uiuc.edu/~miksa/LCAVision.html)

Grand Challenge Cosmology Consortium (http://lca.ncsa.uiuc.edu/GC3Home.html)

Biographical Sketch:

Michael L. Norman

Michael Norman earned a B.S. in astronomy at the California Institute of Technology in 1975, and M.S. and Ph.D. degrees in engineering and applied science from UC Davis in 1976 and 1980, respectively, while a student employee in B Division at LLNL. Working under the supervision of James R. Wilson, he did a thesis in computational astrophysics, a field he has worked in ever since. Norman has worked on a variety of problems in astrophysics and cosmology, including star formation, supernova remnants, astrophysical jets, and most recently, the formation of cosmic structure.

He is currently a professor of physics at the University of California, San Diego, where he directs the Laboratory for Computational Astrophysics. He has previously held positions at the Max Planck Institute for Astrophysics, Los Alamos National Lab, and for 14 years he was a senior research scientist and team leader at the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana-Champaign.

In 1992, Norman founded the Laboratory for Computational Astrophysics (LCA), which develops and disseminates application software for astrophysical fluid dynamical simulations. The ZEUS family of radiation magnetohydrodynamic codes is in worldwide use. He is currently working on numerical algorithms for 3D radiative transfer and radiation hydrodynamics for astrophysical and cosmological applications. While at Illinois, he was also a scientific team leader of the ASCI ASAP Center for the Simulation of Advanced Rockets.

He has received Germany's Alexander von Humboldt Research Prize, the IEEE Sidney Fernbach Award, as well as several Supercomputing'XY HPCC Challenge Awards.

January 24, 2001

Wavelets: An Overview, with Recent Applications

Ingrid Daubechies
Princeton University

Email: ingrid@math.Princeton.edu

Abstract:

Wavelets have emerged in the last decade as a synthesis from many disciplines, ranging from pure mathematics (where forerunners were used to study singular integral operators) to electrical engineering (quadrature mirror filters), borrowing in passing from quantum physics, from geophysics, and from computer-aided design. The first part of the talk will present an overview of the ideas in wavelet theory, in particular wavelet bases. The second part of the talk will discuss some recent applications.

Speaker's web page: http://www.princeton.edu/~icd/

Institution web page: http://www.princeton.edu/

Biographical Sketch:

Ingrid Daubechies

Ingrid Daubechies received her Bachelor's and Ph.D. degrees from the Free University in Brussels, Belgium, in physics, in 1975 and 1980, respectively. She held a research position at the Free University until 1987. From 1987 to 1994 she was a member of the technical staff at AT&T Bell Laboratories, during which time she took leaves to spend six months (in 1990) at the University of Michigan, and two years (1991-93) at Rutgers University. She is now a Professor of Mathematics at Princeton University and is affiliated with Princeton's program in Applied and Computational Mathematics. Her research interests focus on the mathematical aspects of time-frequency analysis, in particular wavelets, as well as applications.

In 1998 she was elected to the National Academy of Sciences and became a Fellow of the Institute of Electrical and Electronics Engineers. The American Mathematical Society awarded her a Leroy P. Steele prize for exposition in 1994 for her book "Ten Lectures on Wavelets," as well as the 1997 Ruth Lyttle Satter prize. Dr. Daubechies was awarded the National Academy of Science Medal in Mathematics in 2000, and the Eduard Rhein Foundation 2000 Basic Research Award for the invention, the mathematical advancement, and the application of wavelets. From 1992 to 1997 she was a fellow of the John D. and Catherine T. MacArthur Foundation. She is a member of the American Academy of Arts and Sciences, the American Mathematical Society, the Mathematical Association of America, the Society for Industrial and Applied Mathematics, and the Institute of Electrical and Electronics Engineers.

November 15, 2000

The Analysis of Algorithms: Experiments versus Theory

David S. Johnson
AT&T Labs

Email: dsj@research.att.com

Abstract:

A common response to a new theoretical result about an algorithm's performance is "Yes, very interesting, but how does it work in practice?" A common response to a paper presenting new experimental results about an algorithmic performance is "Interesting, but do these specific data points tell us anything of more general significance?" Thus, theory and experiment are often placed in opposition. In this talk I will show how they in fact can work together, with illustrations from the study of the "bin packing" problem.

In the simplest version of bin packing, one is given a sequence of items with sizes between 0 and 1, and desires to pack them into a minimum number of unit-capacity bins. This problem has many applications, from optimizing file storage on floppy disks to packing the data for Internet phone calls into ATM packets. The bin packing problem has for three decades served as one of the main test beds for new algorithmic ideas and modes of analysis. From the very beginning, experimental analysis has proceeded in parallel with theoretical work, but in recent years feedback between the two has led to surprising insights and new algorithms of unexpected power.

Speaker's web page: http://www.research.att.com/~dsj/

Institution Web page: http://www.research.att.com/info/Projects/

Biographical Sketch:

David S. Johnson

David S. Johnson was born in Washington, DC, in 1945 and grew up in Milwaukee, Wisconsin. He received a BA summa cum laude in Mathematics from Amherst College in 1967, and a Ph.D. from MIT in Mathematics in 1973, with a thesis entitled "Near-Optimal Bin Packing Algorithms." In 1973 he joined AT&T where he has worked to the present day (although his immediate employer has changed from Bell Labs to AT&T Bell Laboratories to AT&T Labs during the course of AT&T's many divestitures). He currently heads the Algorithms and Optimization Department at AT&T Labs - Research in Florham Park, NJ.

Dr. Johnson is perhaps most widely known for the book *Computers and Intractability: A Guide to the Theory of NP-Completeness*, which he co-authored with Michael Garey and for which they won the 1979 Lanchester Prize of the Operations Research Society of America. NP-complete problems are conjectured to be impossible to solve exactly with provably efficient algorithms, and throughout his research career Dr. Johnson has been interested in how one copes with such theoretical intractability. He was one of the pioneers in the field of performance guarantees for "approximation algorithms," fast algorithms for optimization problems that are not guaranteed to find optimal solutions, but hope to find near-optimal ones. An example of such an algorithm is the "First Fit Decreasing" algorithm for bin packing, which he proved in his Ph.D. thesis never uses more than 11/9 times the optimal number of bins (at least in an asymptotic sense).

In recent years, he has become interested in the average-case analysis of heuristics, and his research has emphasized the interplay between theoretical and experimental analysis in this domain. His experimental work has concentrated on fundamental problem domains such as graph coloring, graph partitioning, bin packing, and the traveling salesman problem, and in the comparison between classical approaches to these problems and new "metaheuristical" approaches such as simulated annealing and genetic algorithms. He has written and talked extensively on methodological issues related to the experimental analysis of algorithms and has overseen the ongoing series of DIMACS Implementation Challenges, which he initiated in 1990.

He has over 100 scientific publications, including 23 editions of a column on the theory of NP-completeness that he writes for the Journal of Algorithms. He is a Fellow of the Association for Computing Machinery (ACM) and is currently a member of the ACM Council and the DIMACS Executive Committee, having served previously as Chair of the ACM Special Interest Group on Algorithms and Computation Theory (SIGACT) and on the Board of the Computing Research Association (CRA).

ISCR

Institute for Scientific Computing Research

University Collaborative Research Program Subcontract Research Summaries

CASC Center for Applied Scientific Computing

Application of Probe-Based Storage to High-Performance Computing

Zachary Peterson and Darrell Long

University of California, Santa Cruz

Summary:

Although there is continual improvement in speed and capacity of storage systems, traditional longitudinal magnetic recording is approaching a hard physical limit. A performance gap between the speed and capacity of RAM and disk is increasing at a rate of 50% a year. This performance gap is especially visible in high-performance workloads such as those employed in NNSA's Advanced Simulation and Computing Program (ASCI).

We in the Computer Systems Research Group (CSRG) at the University of California, Santa Cruz, are investigating two lines of new and innovative high-performance storage research. The first involves an exciting new storage technology based on MicroElectroMechanical Systems (MEMS), which promises a significant increase in performance, capacity, and reliability relative to modern storage devices. The second approaches the performance problems of current storage devices, applying novel ideas of grouping like data to improve read performance, while eliminating the amount of rewriting needed to be done to keep these data contiguous. Using sanitized workload traces from LLNL and support from the ISCR program, we have shown both lines of research successful in addressing the I/O problems posed by ASCI applications.

Recent advances in MEMS technology have given light to exciting new possibilities in storage technology. Using arrays of tiny read/write heads and electrostatic activated microactuators in conjunction with media that moves two dimensions, MEMS-based storage promises huge advantages over contemporary magnetic media. These advantages include: a higher bit density per square inch, a storage of less mass that requires less power to operate, an efficient device-based redundancy of data, and parallelism that increases the data access rate to up to 40 times that of current storage devices. The CSRG has led investigative research in these innovative devices. We have begun to examine the low-level characteristics of MEMS-based storage through the use of simulated device modeling, which enables us to better understand how this new technology can best serve a high-performance computing environment. The CSRG has three device models that address the performance memory gap in different ways: as a replacement for mass secondary storage devices, as a new layer of cache in the memory hierarchy, and as a component of a storage subsystem acting as a large read/write buffer for "hot" metadata and data. Each of these device applications addresses a different aspect of the memory performance gap, meaning all models are feasible in a single high-performance system, providing high throughput and reduced latency.

Beyond accurate and innovative device modeling, we continue our investigation by looking at how I/O scheduling affects device performance. We have been able to not only apply existing policies and algorithms, but also to develop new scheduling algorithms that take advantage of the unique characteristics of this implicitly parallel device. By varying sled movement and the way data is placed in the device, we can better serve I/O requests both satisfying fairness and performance requirements. Our results show that

Summary (continued):

Zachary Peterson and Darrell Long
University of California, Santa Cruz

algorithms developed here in the CSRG outperform almost all algorithms intended for standard magnetic media. It is clear that an application of existing technology to this device is not a sufficient or acceptable solution.

We have been able to make excellent progress on modeling, architectural alternatives, and I/O scheduling. However, there is still research to be done in these areas as well as in aspects of data layout to fully take advantage of the exciting possibilities offered by MEMS-based storage.

Email: zachary@cse.ucsc.edu

Personal web page: http://people.ucsc.edu/~zacharyp/

Institution web page: http://www.ucsc.edu/

Mining Large Image Datasets

Jelena Tesic and Bangalore Manjunath

University of California, Santa Barbara

Summary:

Mining large image datasets places a number of challenging requirements on the analysis framework. Some initial success has been achieved with systems that represent images as an organized collection of summarized information obtained from the feature descriptors and spatial constraints. However, the high dimensionality of the feature spaces and the size of the image datasets make meaningful summarization a challenging problem. A visual thesaurus based on low-level image descriptors provides a scalable conceptual framework for analyzing perceptual events. The heart of this method is a learning system that gathers information by interacting with database users.

Our main objective is to find clusters that represent similar feature points located in a small subset of a feature space. High-dimensional spaces represent challenges for clustering, due to the sparseness of the space. However, clusters may be formed from a couple of visually different elements that inhabit a large part of a high-dimensional space. Co-occurrence of clusters in an image helps us distinguish visually meaningful representatives. We are currently conducting experiments on texture feature sets to determine the dependency of the clusters in the texture feature space on feature vectors and spatial image layout.

The objective of the visual thesaurus is to classify the image regions into perceptually similar categories. Spatial Event Cubes (SECs) are used to represent and analyze the spatial relationships. SECs are computed with respect to particular spatial relationships. Detailed analysis shows that SECs can be used for visualization, discovery of latent spatial configurations, and for constructing indices for efficient and meaningful data access.

Email: jelena@ece.ucsb.edu

Institution website: http://www.ucsb.edu/

Personal website: http://vision.ece.ucsb.edu/~jelena/

Sensitivity and Uncertainty Analysis for Large-Scale Differential-Algebraic Systems

Yang Cao and Linda Petzold

University of California, Santa Barbara

Summary:

Sensitivity analysis generates essential information for design optimization, parameter estimation, optimal control, data assimilation, process sensitivity, and experimental design. In our earlier work, the DASPK3.0 software package was developed for forward sensitivity analysis of differential-algebraic (DAE) systems. Some problems require the sensitivities with respect to a large number of parameters. For these problems, particularly if the number of state variables is also large, the forward sensitivity approach is intractable. We have developed an efficient algorithm for sensitivity computation of large-scale differential-algebraic systems based on the adjoint method. The new algorithm is more efficient than forward sensitivity analysis when there are a large number of parameters. We have analyzed the issues of stability of the adjoint problem and its numerical solution, the determination of consistent boundary conditions for the adjoint system, and the conditions under which the sensitivity analysis problem is well posed. Software called DASPKADJOINT based on (an extended version of) DASPK3.0 and our new adjoint method was developed.

We are finding that the adjoint method is a very powerful tool for the estimation of computational and modeling errors in general, and we have begun work on applying it to the estimation of errors of reduced and/or simplified models. In a surprising twist, we have discovered an adjoint-based condition estimator for a simpler class of problems: solution of linear systems of equations. The new condition estimator appears to be more accurate and efficient than current techniques, and is also adaptable to a greater variety of needs. We have also begun work on both forward and adjoint sensitivity analysis for PDEs. The adjoint problem for PDEs has many challenges, due to the difficulties in finding the adjoint for general problems with different boundary conditions, and also when using nonlinear and adaptive discretization schemes.

Email: petzold@engineering.ucsb.edu

Institutional website: http://www.ucsb.edu/

Personal website: http://www.engineering.ucsb.edu/~cse/

Simulation of Compressible Turbulent Flows With Reaction

David Lopez, Carlos Pantano, and Sutanu Sarkar

University of California, San Diego

Summary:

Two outstanding problems in the area of compressible reactive flows are being investigated. The first problem concerns large eddy simulation where the strongly nonlinear dependence of the reaction rate term on temperature leads to subgrid contribution in the resolved-scale equations, which must be modeled. We have developed a subgrid model by using the information available in the resolved scales along with additional physical input, namely, a model spectrum for the unresolved scales. Promising results have been obtained when evaluating this subgrid reaction rate model against a direct numerical simulation of a shear layer. The second problem occurs when a burn initiated in a NIF capsule encounters inhomogeneities of mixture fraction due to the introduction of inert shell material into DT mix by Rayleigh–Taylor or Richtmeyer–Meshkov instabilities. Preliminary results have been obtained in a simple model problem to identify the important parameters that control the modified burn propagation.

Email: sarkar@mechanics.ucsd.edu

Personal web page: http://www-mae.ucsd.edu/RESEARCH/SARKAR/sarkar.html

Institution web page: http://www.ucsd.edu/

Probabilistic Clustering of Dynamic Trajectories for Scientific Data Mining

Scott Gaffney and Padhraic Smyth

University of California, Irvine

Summary:

Data-driven exploration of massive spatio-temporal data sets is an area where there is particular need of new data analysis and data mining techniques. Analysis of spatio-temporal data is inherently challenging, yet most current research in data mining is focused on algorithms based on more traditional feature-vector data representations. The goal of this research is to develop a flexible and robust framework (as well as algorithms and software tools) for tracking and clustering time-trajectories of coherent structures in spatio-temporal grid data. The benefit of such an approach is to begin to provide a basic set of data analysis tools for exploration and modeling of dynamic objects, in a manner analogous to the much more widely available techniques for clustering of multivariate vector data (e.g., k-means, Gaussian mixtures, hierarchical clustering, etc).

To date in this work we have investigated a general probabilistic approach to clustering of sets of trajectories. We assume that the set of trajectories was generated by a mixture of K component trajectory models, where each component model describes a particular type of trajectory. We have implemented in MATLAB a flexible EM-based clustering algorithm based on these ideas. The algorithm takes as input a set of trajectories, where each is described as a multivariate vector of measurements over time, e.g., (x,y) position and/or various object features measured at each time *t*. The trajectories can be of different time-durations and can have measurements at different times. Different forms of component trajectory model for each cluster can be selected by the user. We have currently implemented the following cluster models: linear regression, polynomial regression, and non-parametric kernel regression. The algorithm is also supplied with a value for *K*, the desired number of clusters. We have also implemented a cross-validation algorithm that selects the best value for *K*, the number of clusters, based on cross-validated predictive probability scores. We have tested the models and the general algorithm on a number of simulated data sets and verified its ability to recover the underlying *K* data-generating trajectory component models from a set of *N* trajectories without labels.

The primary application of trajectory clustering that we have investigated up to this point is the clustering of extra-tropical cyclone (ETC) tracks from meteorological data over the earth's northern hemisphere. ETCs are significant for a number of reasons. For example, they are responsible for severe and highly damaging weather over North America and western Europe. In addition, it is not yet well-understood how ETC patterns are correlated with long-term climatic phenomena and what the implications of global warming might be on ETC frequency, intensity, and spatial distribution. Thus, quantification of the spatial and temporal patterns of ETCs is well motivated.

We have analyzed 15 winters of the mean sea-level pressure field for the CCM3 AMIP II data set, at 6-hourly intervals, on a 2.5-degree square grid over the North Atlantic. Using this data we have developed an algorithm (and implemented it in MATLAB) to locate and track candidate cyclone centers in each 2D frame of the data. The resulting cyclone trajectories were then used as input to our MATLAB model-based clustering algorithm. Analysis of the results suggest that a *K*=3 model with linear regression components provides both a good fit to the data as well as being quite interpretable from an atmospheric science

Summary (Continued):

Scott Gaffney and Padhraic Smyth

University of California, Irvine

viewpoint. We are currently in the process of analyzing the clustering results with atmospheric science collaborators at UCLA (Andy Robertson and Michael Ghil). From a scientific viewpoint there are a number of open questions in terms of interpretation of the cluster models: How should the results be visualized? Can we analyze intensity profiles as well positional information? How predictable are the cyclone trajectories? Are the clusters correlated with other, more global atmospheric phenomena such as the North Atlantic Oscillation index? We are currently in the process of investigating these questions.

On a more general note, we are also investigating how our methodology can be generalized to other types of trajectory data analysis problems. For example, the spatial information (in the form of morphological features) is of interest in many applications and a natural question is how to incorporate this in the clustering. Another question of interest is the sensitivity of our methodology to the particular component models being used: if the trajectories are not truly linear, but linear component models are used, how does this affect the quality of the results? We are investigating the use of statistical random effects models as a methodology to generalize in the direction of a more flexible model.

Email: smyth@ics.uci.edu

Institutional website: http://www.uci.edu/

Personal website: http://www1.ics.uci.edu/~smyth/

Institute for Scientific Computing Research

ISCR Subcontract Research Summaries

Large Scalable Algebraic Domain Decomposition Preconditioners

Randolph E. Bank
University of California, San Diego

Summary:

Many leading-edge scientific and engineering simulations expend large amounts of computational resources for the solution of linear systems of equations. Multilevel and domain decomposition methods have been identified as potentially scalable linear system solvers on terascale computer platforms. However, interprocessor communication overheads, degree of parallelism (in solving the coarse problems), and effects of increasing number of processors on convergence rates all contribute to the list of obstacles to true scalability. The novel feature of our domain decomposition approach is that the subproblem residing in each processor is defined over the entire domain, although the vast majority of unknowns for each subproblem are associated with the subdomain owned by the corresponding processor. The purpose is to ensure that a global coarse description of the problem is contained within each of the subproblems, so a separate coarse grid solve is not required to achieve rapid convergence of the overall iteration. This procedure is similar in philosophy to the parallel adaptive mesh refinement paradigm introduced by Bank and Holst, in that it attempts to both maximize the use of existing sequential algebraic solvers on each processor, and minimize as much as possible the communication between processors. This approach can be applied to general sparse matrices, although matrices arising from discretizations of partial differential equations are the principal target.

In a classic domain decomposition (DD) algorithm, each processor is responsible for solving a linear system corresponding to its subdomain. The DD preconditioner can often be improved by adding overlap, typically allowing processors to solve for certain unknowns near the interface but outside its domain. Additionally, a coarse grid correction step is often applied to capture the global low frequency error modes, a necessary step for rapid convergence. Our philosophy extends these ideas to a natural conclusion; in some sense each processor is given the entire problem to solve, so there is maximal overlap, allowing each processor to independently resolve the global low frequency errors. However, on the parts of the problem it does not own, it uses a coarsened approximation. These subproblems are solved locally on each processor, using any efficient sequential solver and no communication. Our hope is that these solves comprise the majority of work involved in the overall solution process, and the number of outer iterations will be small due to rapid convergence of the global iteration, leading to an efficient scalable algorithm with low communication.

A preliminary version of a DD procedure based on this paradigm was implemented in the finite element package PLTMG, and the computational results are quite promising. It appears from two to ten global iterations are necessary for the class of problems solved by PLTMG (scalar elliptic PDEs), with each global iteration requiring just two communication steps — a boundary exchange near the beginning of each iteration, and a few scalars (norms and inner products used in convergence criteria and damping) near the bottom of the iteration. However, in PLTMG the issue of coarsening does not arise directly. The problems which are refined in a given processor's subdomain and coarse in the remaining regions are provided naturally by the Bank–Holst parallel adaptive meshing paradigm.

Summary (continued):

Randolph E. Bank
University of California, San Diego

Another algorithm based on this approach was proposed by Bank and Jimack and analyzed in a manuscript by Bank, Jimack, Nadeem, and Nepomnyaschikh, where it was shown to be optimal. Here again, the structure of the coarse mesh was provided a priori, and the details of this grid structure played an important role in the theoretical analysis.

The next and major step in the project will be to develop a similar algorithm, basing the implementation as much as possible on existing routines in the *hypre* library at LLNL. Since *hypre* already has excellent algebraic solvers that can be used in this application, the main focus will be on building the coarse approximation for regions outside of the given processor's subdomain. This coarsening problem is related to the construction of algebraic multilevel pre-conditioners, but it is simplified because a hierarchy of levels is not needed. It is hoped that existing routines in *hypre* can be adapted to the requirements of this approach. If successful, this inherently parallel DD method will offer an attractive and easily implemented alternative for solving large sparse systems of equations arising from discretized PDEs.

Multilevel Nonlinear Additive Schwarz Preconditioned Inexact Newton Methods and Applications

Xiao-Chuan Cai
University of Colorado

Summary:

The focus of our research has been on the design, analysis, and software implementation of a new family of multilevel additive Schwarz preconditioned inexact Newton methods for solving nonlinear algebraic systems of equations arising from the finite element or finite difference discretization of nonlinear partial differential equations. Inexact Newton is a very fast algorithm for solving systems whose nonlinearities are well balanced; however, such balance rarely exists in the solution of many practically important problems, such as high Reynolds number incompressible flows or transonic compressible flows. A technique of interest for such cases is to precondition the nonlinear system before calling the inexact Newton solver. This idea has recently been proven effective when using a single-level additive Schwarz method such as the nonlinear preconditioner (ASPIN).

The single level ASPIN is partly scalable in the sense that when the mesh is refined the number of nonlinear outer Newton iterations doesn't change much, and the number of global linear iterations for solving the Jacobian system doesn't change much either. However, for elliptically dominated problems when the number of processors increases, though the number of Newton iterations remains fixed, the number of global linear iterations per Newton iteration increases considerably. This limits the usefulness of the method on ASCI-scale computers. To make ASPIN scalable with respect to the number of processors in a parallel computer, we are currently developing several multilevel versions.

One approach is based on the classical two-level additive Schwarz method in which a coarse solution is simply added to the local solutions. We have showed theoretically that the method provides a nonlinear system that is equivalent to the original system and showed numerically that the number of global linear iterations does not increase much when the number of processors increases.

Our other approach is based on the FAS method, which uses the coarse space correction in a multiplicative way. The complexity inherent in such software for parallel multilevel nested nonlinear iterations demands a computational framework that provides broad-based infrastructure. Our software is written using Argonne National Lab's PETSc, which integrates a hierarchy of components that range from low-level distributed data structures for grids, vectors, and matrices through high-level linear and nonlinear solvers.

Several test cases have been studied using ASPIN. This includes the incompressible Navier–Stokes equation, the compressible full potential equation, and a non-resistive magnetohydrodynamic flow problem.

First-Order System Least Squares for Linear and Nonlinear Elasticity Equations

Zhiqiang Cai
Purdue University

Summary:

Basic equations of elasticity are generally in self-adjoint form, so they lend themselves naturally to an energy minimization principle, cast in terms of the primitive displacement variables. Unfortunately, this direct approach seems to have many practical difficulties (e.g., degraded approximation properties of the discretization and convergence properties of the solution process) as the material tends to become incompressible (i.e., the Poisson ratio tends to 0.5 from below). There have been several attempts to develop alternative approaches that are robust in the incompressible limit. Compounding these difficulties is the fact that what is often needed in practice is the elements of the stress tensor. These variables can be obtained by differentiating displacements, but this weakens the order and strength of the approximation.

The practical need of the stress tensor motivated extensive studies of mixed-finite element methods in the stress-displacement formulation. Unlike mixed methods for second-order scalar elliptic boundary value problems, stress-displacement finite elements are extremely difficult to construct. This is because the stress tensor is symmetric. A beautiful finite element space had not been constructed until recently by Arnold and Winther. Their space is a natural extension of the Raviart–Thomas space of H(div). Previous work imposed the symmetry condition weakly via a Lagrange multiplier. The minimum degree of freedom on each triangle of Arnold–Winther space for the symmetric stress tensor in two dimensions is twenty-four, which is very expensive. Like scalar elliptic problems, mixed methods lead to saddle-point problems and mixed finite elements are subject to the inf-sup condition. Many solution methods that work well for symmetric positive problems cannot be applied directly. Although substantial progress in solution methods for saddle-point problems has been achieved, these problems may still be difficult and expensive to solve.

In the recent years there has been a serious interest in least-squares methods. A number of least-squares formulations have been proposed, analyzed, and implemented. In particular, the least-squares method by Cai, Manteuffel, and McCormick aims to compute the stress tensor directly and, hence, accurately, and it is robust in the incompressible limit. This method is a two-stage algorithm that first solves for the gradients of displacement (which immediately yield stress tensor), then for the displacement itself (if desired). Under certain H2 regularity assumptions, it admits optimal H1-like performance for standard finite element discretization and standard multigrid and domain decomposition solution methods that is uniform in the Poisson ratio for all variables. A limitation of this approach is the requirement of sufficient smoothness of the original problem. Also, the gradient of displacement is not an immediate physical quantity and it is hard to extend this approach to nonlinear elasticity.

With goals of the accurate approximation to the stress, robustness in the incompressible limit, efficient solvers, and applicability to nonlinear elasticity, we developed a least-squares finite element method based on the stress-displacement formulation. As we mentioned before, a major numerical difficulty is how to handle the symmetry of the stress tensor in the stress-displacement formulation. To circumvent such a difficulty, we impose the symmetry condi-

Summary (continued):

Zhiqiang Cai
Purdue University

tion in the first-order system and then apply the least-squares principle to this over-determined, but consistent system. The least-squares functional uses the L2 norm and it is shown that the homogeneous functional is equivalent to the energy norm involving the Lamé constant for the displacement and the standard H(div) norm for the stress. This implies that our least-squares finite element method, using the respective Crouzeix–Raviart and Raviart–Thomas spaces for the displacement and stress, yields optimal error estimates uniform in the incompressible limit. The total number of degrees of freedom is twelve per triangle in two dimensions and eighteen per tetrahedron in three dimensions. This work has been written up for submission to *SIAM J. Numer. Anal.* The algebraic system resulting in this discretization may be efficiently solved by multigrid methods, which is the topic of our current study. We will also continue our effort by the important extension of this approach to nonlinear elasticity and possible applications in the ALE3D project in LLNL.

Investigation of the Richtmyer–Mechkov Instability Using Adaptive Mesh and Algorithm Refinement

Alejandro Garcia
San Jose State University

Summary:

In recent years there has been a growing interest in fluid mechanics at microscopic length and time scales. At these scales the continuum representation of hydrodynamics, as expressed by partial differential equations, breaks down and computational fluid mechanics (CFD), whose numerical schemes are based on such a macroscopic description, is not accurate. For such regimes, algorithms based on statistical mechanics have been developed and shown to capture accurately all the relevant physics. The most well-known examples of such simulation methods are molecular dynamics (MD) and direct simulation Monte Carlo (DSMC); the greatest disadvantage of these molecular simulations is their computational expense relative to CFD methods.

For some problems only a small (but important) volume in a calculation requires molecular resolution. This scenario is reminiscent of adaptive mesh refinement where only a limited region, say near a shock front, requires high resolution. This association led to the development of Adaptive Mesh and Algorithm Refinement (AMAR), which embeds a DSMC simulation within an adaptive mesh refinement framework.

Recently, Rich Hornung (LLNL), Sanith Wijesinghe (MIT), and I have developed an advanced AMAR code using SAMRAI (Structured Adaptive Mesh Refinement Applications Infrastructure). We have also benefited from the assistance of SAMRAI team members (Steve Smith, Andy Wissink, et al.) and Nicholas Hadjiconstantinou, Wijesinghe's doctoral advisor at MIT. Using the SAMRAI framework has allowed us to design a code that has multiple DSMC regions and is fully adaptive; multiple refinement and de-refinement criteria can be imposed. SAMRAI-based communications routines handle the passing of particle information between patches (e.g., copying particles from one DSMC patch into the ghost cells of another). The hierarchical data in the simulation is output in the form of plot files formatted for the Vizamrai suite of visualization tools.

The first application for which the AMAR/SAMRAI code was originally targeted is the Richtmyer–Meshkov instability, that is, the enhancement of mixing in a binary gas due to the impulsive acceleration from a shock wave. This past year multi-species functionality was added to the Euler solver. A variety of test cases have been designed to test these new routines and their coupling with the DSMC portion of the program. Until recently it was difficult to obtain quantitative results from the Vizamrai-based output files, but this summer Steve Smith provided us with new data reduction routines to "crunch" these hierarchical data files into a simplified form that can be processed by standard numerical tools such as Matlab. The analysis of this data, as well as further validation testing, is currently underway.

Despite the improved efficiency of a hybrid code versus a conventional particle simulation, our problems of interest require large-scale computations best performed on a massively parallel computer. Our work this summer showed that the heavy communications and memory demands of a particle method require the development of more sophisticated load balancing strategies. The potentially high cost of particle data redistribution during the adaptive gridding phases of the solution algorithm should also be investigated. Recently the memory management in the DSMC routines was redesigned, which we hope will allow us to run significantly larger problems. Parallel algorithm questions such as these are of increasing interest in computational science research, especially for multi-scale and multi-algorithm methods such as AMAR.

Lagrange–Newton–Krylov–Schur Solvers for PDE-Constrained Optimization

Omar Ghattas
Carnegie Mellon University

Summary:

This project focuses on Lagrange–Newton–Krylov–Schur (LNKS) algorithms for large-scale optimization problems that are constrained by systems of partial differential equations and their applications to optimal control, optimal design, and parameter estimation problems in science and engineering.

Collectively, we refer to these as "inverse problems" to distinguish them from "forward problems" that usually characterize large-scale simulations. In the forward problem, problem data—initial conditions, boundary conditions, material coefficients, and the domain geometry—are specified, and the state of the system is found by solving the PDEs. The inverse problem involves the reverse process: some components of the state are typically specified (through an objective function to be minimized), and solution of the PDE-constrained optimization problem yields components of the data, often called the decision variables.

The inverse problem is often significantly more difficult to solve than the forward problem. This is because the inverse problem includes the PDEs as part of its constraint set, and because the inverse problem is often ill-posed despite the well-posedness of the forward problem. Not surprisingly, most of the work in large-scale simulation has been directed at the forward problem. Sustained advances over the past twenty years have produced a body of efficient parallel scalable algorithms for many classes of PDE simulations. This invites research into what is often the ultimate goal in many areas of computational science and engineering: the optimal design, optimal control, or parameter estimation problem, in the form of a PDE-constrained optimization problem.

During my sabbatical at LLNL, I worked on LNKS algorithms for time-dependent PDE-constrained optimization in conjunction with CMU graduate student Volkan Akcelik. One of the difficulties with time-dependent problems is that optimization "sees" the entire time history, and therefore the optimization problem is coupled across time. We tailored Newton–Krylov-type methods for time domain inverse wave propagation problems, which serve as a good testbed for time-dependent optimization. The specific class of problems studied was estimation of material parameters in an acoustic medium, given waveform observations at specified locations on the boundary. The Newton–Krylov optimization methods were implemented on top of the parallel PDE solver library PETSc. Problems of up to several hundred thousand inversion parameters were solved, on up to 128 T3E processors. The mesh-independence of the Newton and Krylov iterations demonstrated algorithmic scalability with respect to number of inversion and state parameters. We also studied issues related to total variation regularization, and to overcoming the multiple minima problem via multiscale extension of the Newton–Krylov solver.

In addition to conducting research on algorithms for time-dependent inversion, during my stay at LLNL I taught a ten-lecture short course entitled "Computational Optimization." The focus of the course was on numerical algorithms for large-scale continuous optimization problems and their particularization for optimization problems that are constrained by discretized

Summary (continued):

Omar Ghattas
Carnegie Mellon University

PDEs. Sensitivity analysis and its relationship to reduced space optimization methods were also discussed. Course participants came from CASC and several science and engineering divisions.

During my stay I also participated in the Nonlinear Solvers and Differential Equations project, and I interacted with the sensitivity analysis group. CASC collaborator Carol Woodward and I worked on formulating optimal control and inversion problems for two problems in environmental and atmospheric sciences that are of interest to LLNL researchers.

A Fast Poisson Solver for General 3D Regions

Anne Greenbaum
University of Washington

Summary:

We are developing a software package for solving Poisson's equation on general three-dimensional regions, in approximately twice the time required for a fast Poisson solver on an embedding cube. The surface of the region must be triangulated, but the interior need never be meshed. The algorithm involves two steps, the first of which consists of evaluating a volume integral and the second of which requires solving Laplace's equation with given boundary conditions. Each can be expressed as a Poisson problem on an embedding cube, provided the discontinuities in the desired solution and its derivatives across the boundary of the region can be determined. This requires solving a second Fredholm boundary integral equation. Such equations are very amenable to solution by iterative methods such as the GMRES algorithm. If there are m boundary discretization points, then the work to solve the integral equation using the GMRES algorithm with dense matrix-vector multiplication is O(m**2), but this can be reduced to O(m) work if the matrix-vector multiplications are carried out with the Fast Multipole Method. The remaining work is that of two Poisson solves on a regular lattice throughout an embedding cube. If there are N lattice points, this is O(N log N) work using an FFT and can even be carried out with O(N) work using multigrid methods (though the constant may be larger). We are writing the code in a modular way so that different algorithms can be substituted for the individual pieces to see which performs best.

The algorithm is amenable to parallelization, and we have tested parts of it on an IBM SP2. The goal was to solve problems on a lattice of (2N)**3 points using 8p processors in about the same amount of time required to solve a problem of size N**3 on p processors—i.e., to have a scalable algorithm. With the fast Poisson solver, the total work is O(N log N) instead of O(N), so we did not expect perfect scalability, but since log N is such a small factor, we did see near perfect scalability. This was very encouraging and shows that the algorithm parallelizes efficiently. We plan to incorporate this package in the LLNL Djehuty code for stellar evolution. Another advantage of this approach, which will be of particular use in the Djehuty code, is that derivatives can be computed directly. This avoids numerical difficulties associated with computing the solution and then using a finite difference formula for its derivatives. Other advantages of this approach include the fact that it can handle exterior problems as easily as interior ones. As mentioned earlier, one never meshes the domain but only the boundary. Additionally, some other equations such as the biharmonic equation, which might prove more difficult for finite difference or finite element techniques, can be solved with this approach almost as easily as Poisson's equation.

Some Aspects on Algebraic Multilevel and Multigrid Methods

Johannes K. Kraus
University of Leoben, Austria

Summary:

Algebraic multilevel and multigrid methods have gained rising popularity in the solution of large sparse systems of linear equations. This is because they often combine the efficiency of classical (structured) multigrid methods with the applicability to unstructured-grid problems. In many situations algebraic multilevel preconditioners, constructed in the framework of incomplete LU decomposition, provide an alternative to classical algebraic multigrid (AMG) methods.

In the first phase of this research project at CASC/LLNL, which was supported by the Austrian Science Foundation (FWF), AMG methods were further examined. The main focus was on a comparative analysis of linear and nonlinear algebraic multilevel iteration (AMLI) based on special hierarchical ordering strategies for the unknowns. In both cases the same assumption on the approximation, caused by neglecting fill-in terms during the LU decomposition, was used. Theoretical and experimental investigations showed that the resulting variable-step preconditioners become very close to linear mappings if we employ approximations to the Schur complements that are close enough on all levels of cyclic reduction. In this case very few, e.g., two or three, inner GCG-type iterations on certain levels, e.g., every other level, are sufficient to achieve optimal order of computational complexity. It is also possible to show similar convergence rates as for the conjugate gradient method accelerated by linear AMLI. Compared to the W-cycle variant of linear AMLI the nonlinear method has the advantage of being free of any method parameters to be estimated. The main results of this work are documented in the paper "An Algebraic Preconditioning Method for M-matrices: Linear Versus Nonlinear Multilevel Iteration," which has been accepted in *Linear Algebra with Applications*.

Another research goal in this project was to investigate element-free interpolation in element-based algebraic multigrid (AMGe) methods. In joint work with Van Emden Henson and Panayot Vassilevski, a new algorithm for constructing neighborhood matrices to be used for the computation of interpolation weights was developed. The method utilizes the existence of simple interpolation matrices (piecewise constant for example) on a hierarchy of coarse spaces (grids). Then one constructs by algebraic means graded away coarse spaces for any given fine-grid neighborhood. Next, the corresponding stiffness matrix is computed on this graded away mesh, and the actual neighborhood matrix is simply the Schur complement of this matrix where degrees of freedom outside the neighborhood have been eliminated. Once the neighborhood matrices have been computed they can serve as "element matrices" in AMGe methods. A description of the algorithm with model complexity analysis as well as some comparative tests of the quality of the resulting improved interpolation matrices is to be submitted for publication ("Computing Interpolation Weights in AMG Based on Multilevel Schur Complements").

Further research was done on new multilevel incomplete factorization algorithms that result in better convergence properties when applied to matrices that are not M-matrices.

Non-Conforming Finite Elements, Mesh Generation, Adaptivity, and Related Solution Methods

Raytcho Lazarov and Joseph Pasciak

Texas A&M University

Summary:

We studied the problem of developing algebraic multilevel algorithms in a parallel computing environment. Specifically, we started with a partitioning of the original domain into subdomains with a generally unstructured finite element mesh on each subdomain. The meshes do not need to be aligned across the subdomain boundaries. An element agglomeration-based algebraic multilevel coarsening is then applied independently in each subdomain. Note that even if one starts with a conforming line grid, independent coarsening generally leads to non-matching grids on the coarser levels. The element-based coarsening gives rise to a face decomposition of each subdomain boundary such that every coarse face is a union of fine faces. We developed a general dual basis mortar approach to set up global problems on all levels based on the above coarsening. The resulting "algebraic" mortar approach is well-defined and consistent in the sense that the generalized dual basis functions reproduce constants locally.

As usual, continuity across the interfaces is imposed by integral constraints against the mortar functions and reduces to a standard dual basis approach when applied to conventional finite element spaces, e.g., the fine level discretization. A general code was developed to illustrate the behavior of the proposed method. This code is based on a number of public domain tools and tools under current development at LLNL. One starts with a conforming coarse grid which is then partitioned into sub-domains using METIS, and the subdomain meshes are refined independently in parallel. An element based agglomeration procedure (AMGe) is then applied to the fine grid problems locally on the subdomains. This AMGe technique operates on and produces (on the coarser levels) generalized "elements" defined in terms of relation tables between elements—faces and degrees of freedom and the local element matrices. This information carries over to the interfaces and enables the local face based construction of the mortar.

Towards Bridging the Processor/ Memory Performance Gap for SMP Systems

Sally A. McKee
University of Utah

Frank Mueller
North Carolina State University

Summary:

Processor speeds are increasing much faster than memory speeds, and this disparity prevents many applications from making effective use of the tremendous computing power of modern microprocessors. Current access times of 50 cycles or more often cause memory performance to dominate application run time, and the processor/memory performance gap continues to grow. The problem is acute for uniprocessor machines, and even worse for symmetric multiprocessors sharing memory resources.

Together with members of CASC, we are developing tools and measures to better understand the memory performance of large, scientific applications. We are then leveraging these to design more efficient memory systems that implement Dynamic Access Optimization (DAO) techniques, which attempt to make better use of the memory system by changing the order and/or apparent location of memory references. This work is in collaboration with SRC Computers, Inc., whose reconfigurable SRC-6 SMP machine serves as the experimental vehicle for our investigations of DAO techniques.

We began by targeting streamed computations with strided access patterns. The principle of locality has guided the design of many key architectural features, including cache hierarchies and TLBs, and quantitative measures of spatial and temporal locality of reference have been useful for predicting the performance of memory hierarchy components. Unfortunately, the concept of locality is constrained to capturing memory access patterns characterized by proximity, while sophisticated memory systems are capable of exploiting other predictable access patterns. For instance, the vectors used in streamed applications lack temporal and often spatial locality, and thus have poor cache behavior. Nonetheless, their access patterns have the advantage of being predictable, and this predictability can be exploited to improve the efficiency of the memory subsystem (e.g., by reordering accesses to avoid bank conflicts or by prefetching stream data within the memory controller).

We address this in part by defining the concepts of spatial and temporal regularity, and by introducing a measure of spatial access regularity to quantify some of the predictability in access patterns. We developed an efficient, online algorithm to dynamically determine the spatial access regularity in an application's memory references, and we have demonstrated its use on a set of regular and irregular codes. We found that the use of our algorithm, with its associated overhead of trace generation, slows typical applications by at least an order of magnitude less than traditional, full-trace generation approaches. Our approach can be applied to the characterization of program access patterns and in the implementation of sophisticated, software-assisted prefetching mechanisms, and its inherently parallel nature makes it well suited for use with multi-threaded programs. We have incorporated this technology into a dynamic instrumentation framework so that we can selectively generate partial data traces through dynamic binary rewriting. This part of our work has produced a workshop paper, a poster presentation, and a conference paper submission.

We are also developing memory microbenchmarks that use hardware performance monitors to measure and categorize latencies in SMP machines. We

Summary (continued):

Sally A. McKee
University of Utah

Frank Mueller
North Carolina State University

have used these benchmarks to evaluate the memory performance of two of the LLNL machines, the results of which were presented at the IBM SP Scientific Computing User Group Meeting (SCICOMP 4).

Current work expands on these efforts and investigates memory-controller architectures and the accompanying software to implement Dynamic Access Optimizations that exploit regularity and other kinds of predictability to deliver better memory system performance.

Discretization of Neutron Transport and Automatic Parallelization

Christoph Pflaum
University of Wurzburg

Summary:

My research at CASC was directed towards two projects. The first project was a discretization of the neutron transport equation. A general problem in the discretization of neutron transport is the approximation of the function space on the sphere. It is important to find an approximation that minimizes so-called "ray effects" and that leads to a simple stiffness matrix.

In cooperation with Britton Chang and Peter Brown, we studied the finite element discretization of a function space on the sphere. The corresponding finite element space consists of piecewise constant functions. Different types of tessellation of the sphere were studied. To obtain a discretization with small ray effects, it is necessary to do a very accurate integration of the finite element functions. Numerical results show that the resulting discretization leads to smaller ray effects than the ordinate direction approximation of the sphere. Furthermore, the resulting stiffness matrix can be evaluated in a very efficient way.

The second project was on automatic parallelization with expression templates on semi-unstructured grids. Manual parallelization of a numerical algorithm tends to be difficult and time consuming. Therefore, one is interested in concepts that lead to an automatic parallelization of algorithms.

An important problem of automatic parallelization is the description of algorithms in a suitable, expressive high-level language. In case of numerical algorithms for PDEs, one would like to implement algorithms in a language that is close to the mathematical language. This can be obtained by expression templates in C++. This concept was used in the library EXPDE for the finite element approximation of PDEs. We implemented a parallel version of this library. The new parallel version of EXPDE contains a parallel grid generator for general domains in 3D, multigrid operators, and several other concepts. The automatic parallelization concept of EXPDE was applied to several PDEs, including the Stokes equations, transport equations, and the equations of elasticity. In cooperation with Rob Falgout, we developed a parallelization concept based on octrees. An important property of this parallelization concept is that the Gauss-Seidel iteration can efficiently be parallelized even in case of relatively small grids. Numerical results on ASCI Blue-Pacific showed the scalability of the parallelization concept for more than 500 processors.

It is planned to apply a finite element approximation of functions on the sphere in the ADRA project. An extension of EXPDE an adaptive grids is also planned.

AMG, Spectral AMGe, and FOSPACK

John Ruge
Front Range Scientific Computations, Inc.

Summary:

Our project focused on AMG, Spectral AMGe, and FOSPACK. Algebraic Multigrid (AMG) is actually a class of methods designed to solve a given discrete problem (generally arising from the discretization of partial differential equations) using multigrid principles by automatically choosing coarser grids, grid transfer operators, and coarser grid operators needed in multigrid cycling. In the most general case, this is done using only the information contained in the matrix itself. This has been shown to be very robust for elliptic problems and can be applied on unstructured meshes and problems with varying or even discontinuous coefficients.

Part of this project involved the examination of fill-in on coarser level matrices while using "classical" AMG codes. While overall complexity appears bounded, stencil size on coarser levels can grow, especially in large 3D problems. This can be an impediment to effective parallelization of the method. Studies on model problems showed that this effect is not due to irregular coarsening near boundaries, as had been hypothesized, since it also appeared when periodic boundary conditions were used on uniform hexahedral meshes. Some growth seems inevitable, even with uniform coarsening patterns. Interpolation truncation was only marginally effective in controlling this growth. Future work here will focus on sparser coarse grids coupled with "long range" interpolation as a treatment for this problem.

Spectral AMGe is an offshoot of classical AMG in which interpolation is constructed in an effort to more directly approximate the lower part of the spectrum of the fine grid operator, since these are the components not effectively reduced by relaxation. This approach, applicable only to finite element discretizations, requires the local stiffness matrices corresponding to the fine grid elements. An agglomeration procedure is used to partition the mesh into groups of elements, and the local matrices corresponding to each group are assembled. The eigenvectors corresponding to the small eigenvalues of these matrices are used to form the columns of interpolation matrix. Such an approach is potentially much more robust than standard AMG approaches. Tests were performed on a FOSLS formulation (see below) of 2D Helmholtz problems, verifying that the method performs well, even in the presence of large near-null spaces of the operator.

The majority of work in this project was devoted to the further development and testing of the FOSPACK code, focusing on problems arising in the ALE3D package. First-order system least squares (FOSLS) is a relatively recent approach to problem formulation, in which a PDE or system of PDEs is reformulated as a system of first-order equations, and the problem is posed as the minimization of the sum of the square of the L2 norms of the corresponding residuals. For many problems, formulations have been obtained in which it can be shown that this functional is equivalent to the square of the H1 norm of the error. One very nice consequence of this is that the resulting system has elliptic diagonal blocks that dominate the off-diagonal blocks. This makes AMG a very attractive solver for such problems. FOSPACK is a program that combines easy problem specification, assembly, and solution of FOSLS formulations. We applied the code to elasticity problems from ALE3D, demonstrating O(h**2) convergence in both displacements and stress, and began development of a suite of diagnostic tools to help the user in the formulation of the FOSLS functional in linear and nonlinear problems.

Development of Object-Oriented Tools for the Numerical Solution of Reactive Flow

Donald W. Schwendeman
Rensselaer Polytechnic Institute

Summary:

Our work involved the development of object-oriented application software, within the Overture framework of codes, for the numerical simulation of high-speed reactive flow. The mathematical model on which the software is based is the reactive Euler equations. The numerical implementation of this model is fairly general and allows for multiple reacting species and reaction rates and a general equation of state with the aim of being able to simulate experimentally observed phenomena in gas or solid explosives. The software is part of the OverBlown package of fluids codes, developed and maintained by Bill Henshaw and the Overture team at CASC. It uses overlapping grids in order to handle general domains and the A++/P++ array class library (developed by Dan Quinlan and the Overture team), which allows parallel processing. The software includes a patch grid-type adaptive mesh refinement (AMR) scheme. The software has been carefully tested for accuracy using existing codes written previously by the author, and it has been used to model the evolution to detonation of reactive samples subject to various initial conditions and within various confinement geometries.

The main work involved the continued development of software for the numerical solution of the reactive Euler equations. The equations describe the nonlinear convection of mixture-averaged hydrodynamic variables, such as density, velocity, pressure and temperature, and the convection and production of a set of reacting species that describe the mixture. The chemistry is modeled by prescribed reaction rates for each component and these rates are highly state sensitive for the problems of interest. An equation of state for the mixture is assumed and involves the heats of reaction for each of the reacting species. For a particular simulation, the equations are to be solved on a given domain subject to various initial conditions and boundary conditions.

A numerical method of solution was implemented for the reactive Euler equations. Within the Overture framework, the numerical method discretizes the equations on a set of overlapping, structured grids that cover the domain of interest. The software package Ogen generates the overlapping grid and provides geometric mapping information for each curvilinear component grid and information concerning the communication of the solution between grids in the overlap region. This information was considered to be given and thus the main task was to implement a numerical method for an individual component grid and then let the existing software, OverBlown, handle the surrounding numerical details (such as interpolation between component grids, the application of boundary conditions, time stepping, graphical interface, and more).

The numerical method chosen was a shock-capturing, Godunov-type scheme. The main new element to the software is the implementation of an AMR scheme. This is a patch-based scheme. Multiple levels of finer AMR grids are built for each base grid and are used to resolve the fine-scale structure near the reaction zone. The AMR grids communicate at the boundaries with their coarser parent grid, or with sister grids at the same level. They may also communicate across the overlap. The refinement of the grid is based on an esti-

Summary (continued):

Donald W. Schwendeman
Rensselaer Polytechnic Institute

mate of spatial and temporal gradients of the numerical solution. This is done by monitoring the magnitude of first and second differences of the solution and the magnitude of the reaction rate. The user has considerable freedom to specify and tune the AMR parameters for a particular problem.

For certain problems it is desirable to use the grid generator to introduce front-tracking grids on the base level in order to reduce the number of AMR grids needed to track and resolve a detonation wave. It may also be desirable from an accuracy point of view to confine a detonation wave to a single component grid at the base level in order to avoid the propagation of a detonation wave along a grid overlap. In order to test this numerical approach, we have modified the software to allow grid insertion and deletion. At present this can be done manually, but work is ongoing to test various numerical strategies for managing the tracking grids automatically based on the solution behavior. For example, it is possible to determine the approximate location of a detonation wave based on the behavior of the reacting species.

A number of problems involving high-speed reactive flow have been identified and are under current investigation using the new software. These problems involve paths to detonation of reactive samples at critical conditions, multidimensional detonation propagation, and detonation failure.

Algebraic Multigrid in an Industrial Environment

Dr. Klaus Stüben

Fraunhofer-Institute for Algorithms and Scientific Computing (FhG-SCAI)

Summary:

We have been working on the further development of a general framework for the application of algebraic multigrid (AMG) to systems of partial differential equations (PDEs). This framework is based on the idea of using an auxiliary matrix—called the "primary" matrix—for coarsening and/or for interpolation in the AMG context. A reasonable primary matrix describes the (pointwise) connectivity structure of some auxiliary "primary" unknown which should, in some sense, represent the connectivity structure for all "real" unknowns in the given system of PDEs and, thus, can be used for pointwise coarsening all unknowns simultaneously.

In practice, there are various ways to apply this framework to solve a given system of PDEs. For instance, if the geometric positions of the mesh points are known to AMG, the coefficients of the primary matrix may be based on geometric distances between points. Alternatively, it may be based on a suitable norm of neighboring blocks of unknowns. While, in the first case, coarsening would be closely related to geometric coarsening, it would correspond to a block-coarsening in the second case. One can also imagine that the primary matrix is defined based on some natural physical quantity for which there is no reasonable equation contained in the original system of PDEs. An example of such a situation would be the pressure in the context of the Navier–Stokes equations.

In addition, there are various ways to construct interpolation based on the resulting coarse levels. In particular, interpolation may be different for each physical unknown (e.g., based on the matrix rows in the original system), it may be the same for each unknown (e.g., based on the matrix rows of the primary matrix), or it may be blockwise.

For the framework to be as flexible as possible, a reasonable primary matrix is either defined internally to AMG or it may be user-provided by rendering the original PDE system by a primary matrix that is fully decoupled from the rest of the system. The latter option makes particular sense in situations where AMG cannot construct a reasonable primary matrix based solely on algebraic information. In many such cases, the user of AMG will be able to define a reasonable matrix based on the underlying physics of the given problem.

Projects in Computational Ecology

Tonya Kostova Vassilevska
Bulgarian Academy of Sciences

Summary:

We focus on the application of mathematical and computational methods to problems from biology. The analytical studies of small, specific ODE systems were the goals of two collaborations, which resulted in two papers on the existence of periodic solutions of some equations with applications in ecology and on the qualitative behavior and bifurcation phenomena of the FitzHugh–Nagumo system.

In addition, collaboration with the Environmental Restoration Division on the effect of oil-production sites on the ecology and biodiversity of habitats has been started. A conceptual object-oriented model has been formulated which will be implemented in C++ software. The model will use a geographic information system as a database. The data will be used to simulate the development of a foodweb consisting of vegetation, rodents, and their predators on a large-scale spatial grid in real time. Further, we shall simulate the appearance of spills and their effect on the foodweb.

A Robust Multilevel Nonlinear Method

Irah Yavneh
Technion Institute

Summary:

In a new collaboration during the summer, mainly with Carol Woodward, Panayot Vassilevski, and Jim Jones of CASC, we began investigating a novel approach for multigrid algorithms for nonlinear problems. There are two "classical" approaches for solving discretized nonlinear partial differential equations by multigrid methods. One is to perform a global linearization (GL), usually by Newton's method or some inexact variant thereof, and solve the resulting linear system approximately by a linear multigrid algorithm; this is then repeated iteratively. The second approach (FAS, by A. Brandt, or NLMG, by W. Hackbusch) is to only perform a local linearization (LL) in the error-smoothing (relaxation) process. Convergence acceleration is then provided by nonlinear coarse-grid operators.

For "nice" problems, both approaches work well and the difference in efficiency is usually not large. But for more difficult problems, the two approaches often exhibit distinct behavior, with GL holding an advantage in some situations and LL in others. We were therefore motivated to develop a method which will be at least as good as the more suitable of these two approaches, and often better than both.

Multigrid methods for nonlinear problems have been studied very extensively. For many problems, robust multigrid methods (such as "Algebraic" and "Black Box" multigrid) are well-known to be effective even in the presence of discontinuous coefficients and domain boundaries. This represents an advantage for GL methods, where the problem solved per iteration is linear. LL approaches, on the other hand, generally require "direct" rediscretization of the nonlinear operator on the coarse grids, which is less robust. However, the attraction basin of fast GL methods is usually small, which means that slow global search methods must often be applied before the GL method becomes effective. This may represent a significant disadvantage.

In a new approach, called a Multilevel Nonlinear Method (MNM), we eliminate the compromise between global and local linearization, by splitting the nonlinear operator into two parts. One is a relatively large linear part (normally obtained by Newton linearization), and the second is the small nonlinear part, which remains after the linearized operator is subtracted off. We use the non-robust FAS-like coarse-grid approximation only for the nonlinear part, while the linear part is approximated by a robust Galerkin coarse-grid operator. We thus gain the advantages of both classical approaches: a fast asymptotic convergence rate with a large attraction basin. The additional computational cost and storage is only a fraction of that of the standard algorithms, partly because the extra work is only performed on the non-finest grid. An adaptive MNM approach and a parallel version have also been conceived.

So far we have only tested the new method on one-dimensional problems, albeit some realistically difficult ones. In particular, we experimented with a nonlinear diffusion problem modeling soil water retention. The results are very encouraging, with the new method performing as predicted by our (so far simplified) analysis. Further developments, both analytical and experimental, are in progress.

ISCR

Institute for Scientific Computing Research

Laboratory Directed Research and Development Project Research Summaries

CASC Center for Applied Scientific Computing

SAVAnTS: Scalable Algorithms for Visualization and Analysis of Terascale Science

Mark Duchaineau

Center for Applied Scientific Computing

Summary:

Wavelets and other hierarchies on large-scale scientific geometry form critically needed computational infrastructure to reduce the overwhelming data sizes and to make possible highly interactive exploration of scientific results. Several difficult challenges must be overcome to create such multi-resolution compression and display capabilities. First, appropriate grids must be devised on highly convoluted scientific surfaces, for which no methods have previously been known. Second, novel wavelet transforms are required that are completely local in their mathematical operations, so as to operate effectively on arbitrarily large data sets. Finally, display algorithms must be devised that read exactly the portion of the compressed data required for any given moment of interaction, and use this to feed graphics hardware and optimize its speed and accuracy. Overall, the goal is to achieve orders of magnitude improvements in storage sizes and interaction rates over current best practice.

During a given run, multi-physics simulation codes on LLNL supercomputers produce dozens of terabytes of data that form a vital component of the Stockpile Stewardship Management Program and other programs. Great strides are being made to increase the efficiency and accuracy of the codes by harnessing tens of thousands of processors using scalable algorithms. However, the efficient and accurate post-computation data handling and interactive exploration must also scale efficiently to reach LLNL's goal of a capability for productive 100 Tflop/s or greater simulation. This project introduces multi-resolution methods to data exploration activities that are especially critical to the Laboratory missions: the compact storage and fast display of variables on material boundaries, orthogonal cut planes of 3D field data, contour surfaces, and transparent volume renderings.

Our focus is on devising new wavelet transforms and other hierarchies, and applying them for compression and accelerated display of volumetric field data, material boundaries, and contour surfaces. We find that the useful information content of a 3D field such as pressure or density tends to be quite sparse. Because wavelets automatically find and exploit coherence in both space/time and frequency/scale, this sparse information content is readily compressed after the application of the appropriate wavelet transform.

Wavelets are well understood for regularly spaced grids filling a 3D block stored in a single computer's memory. However, innovations are required for large-scale Laboratory applications, including:

(1) highly adaptive or unstructured settings,

(2) arbitrary surfaces that typically can not be represented as anything resembling a regularly spaced grid, and

(3) data distributed over ten thousand processors.

Fundamentally, the work performed should be proportional to the sparse post-transform information content at as many stages as possible of the end-to-end data flow going from simulation to scientist. This leads to a suite of connected optimization problems that we have addressed.

Summary (continued):

Mark Duchaineau
Center for Applied Scientific Computing

During FY2001 we devised:

(1) a hierarchy-building process for surfaces, extended to exploit time coherence;

(2) new fast and accurate algorithms to extract material boundaries from volume fraction information and produce a related 3D fieldhierarchy;

(3) a shink-wrap surface remapping method that optimizes subsequent display-hierarchy accuracy;

(4) a volumetric-based surface compressor that overcomes the complex-topology limitations of explicit surface mappings;

(5) a spatial index remapper that automatically optimizes cache coherence and enables out-of-core computation on huge data sets;

(6) a magnifying-lens capability for hierarchical hardware-based transparent volume rendering;

(7) a wavelet compression technique that allows contour topology reservation;

(8) a memory-insensitive technique for surface simplification;

(9) a highly simplified and effective height-map display accelerator.

Several large-scale simulations were performed in the fall of 2000 on several thousand processors of the initial-delivery ASCI White supercomputer at LLNL. In collaboration with Farid Abraham of IBM Almaden Research, we have successfully studied for the first time the supersonic propagation of cracks and the formation of complex junction structures in metals. These unprecedented computations were made possible by our project's prototype software for wavelet compression and hierarchical display optimization. In one simulation on 5120 processors, it was demonstrated that it is possible to reduce the anticipated 25 terabytes of output to under a terabyte without perceptible degradation in visualization results, and that the compression work added only 10% to the simulation run time.

Overcoming the Memory Wall in SMP-Based Systems

Bronis R. de Supinski, Andy Yoo, Sally A. McKee, Frank Mueller, and Tushar Mohan

Center for Applied Scientific Computing

Summary:

Both CPU and memory speeds are increasing at exponential rates, as expressed in Moore's Law. Unfortunately, memory hardware is currently significantly slower than CPUs. For example, on snow, the ASCI White testbed system, an average of 87 floating point operations can be completed in the time required to load one operand from main memory. Even worse, CPU speeds are increasing faster than memory speeds; thus, the number of CPU cycles required to access memory is increasing. This divergence will exacerbate an existing problem for codes with large memory footprints, including the codes typically in use at LLNL: memory accesses dominate performance. Not only is the performance of many LLNL codes dominated by the cost of main memory accesses, but many current trends in computer architecture will lead to substantial degradation of the percentage of peak performance obtained by these codes. Many researchers anticipate a "Memory Wall" in which memory accesses imply an absolute performance limit, and improvements in CPU speed provide no performance benefit.

This project is extending dynamic access optimizations (DAO), a promising set of techniques for overcoming the Memory Wall, to symmetric multiprocessors (SMPs), which are common at LLNL. We expect future computer systems to become available to LLNL that use the novel techniques that we are designing to alleviate this problem in SMP-based systems. Further, our techniques complement other emerging mechanisms for improving memory system performance that will be the basis of future LLNL systems, such as processors-in-memory.

DAO techniques have shown significant promise to overcome the Memory Wall without requiring complex source code changes. These techniques change the order or apparent locations of memory accesses from those generated by the issuing program to ones that use the memory system more effectively without changing the results. For example, altering the execution order can exploit memory hardware characteristics such as interleaved memory banks and hot dynamic random access memory (DRAM) rows, while techniques that alter the apparent location can significantly increase cache hit ratios. DAO mechanisms can reduce run times of memory intensive portions of programs by factors of two to an order of magnitude. Previous projects investigating DAO focus on uniprocessor systems and require special-purpose hardware. Although promising, DAO techniques for uniprocessors do not target the systems in use at LLNL. All major LLNL computing resources are clusters with SMP nodes. Thus, we need DAO techniques that support simultaneous access to the memory system by multiple processors. Difficulties arise in SMPs for both types of DAO techniques, access reordering, and shadow memory. This project addresses these difficulties.

During FY01, the first year of the project, we established the infrastructure required for our research. The overall plan is not only to design DAO techniques for SMP-based systems, but to demonstrate that they significantly improve performance for typical LLNL codes. In order to achieve this, we will predict the benefits of these techniques with an analytic model that we are designing, and we will simulate the variety of techniques that we design, as well as implementing the most promising techniques on an actual SMP.

Summary (continued):

Bronis R. de Supinski, Andy Yoo, Sally A. McKee, Frank Mueller, and Tushar Mohan

Center for Applied Scientific Computing

In FY01, we have implemented a simulation that will serve as the basis of all of our simulation studies. In addition, we have gathered the input data of our analytic model for UMT, an LLNL 3D neutral particle transport code for unstructured meshes. This data includes both a characterization of the access patterns of UMT and detailed measurements of the memory systems of several SMP-based systems. All of this data was gathered with new tools that we implemented as part of this project.

An important aspect of our project is that we will not only design SMP-aware DAO techniques and demonstrate their efficacy through simulation, but we will also implement the techniques on an actual system and demonstrate their benefit for UMT on that system. Thus, two of our FY01 milestones involved the selection of an experimental system and an evaluation on the hardware we would use to implement our DAO techniques. After the project started, we established a close collaboration with SRC Computers of Colorado Springs, which includes privileged use of their new SMP. Since its memory controllers, cache coherence controllers, network bridges, and switches are all implemented with field programmable gate arrays (FPGAs), we can implement our SMP-aware DAO techniques directly in the hardware by changing the FPGA programs of these devices.

In FY02, the second year of this project, we will simulate and implement the first SMP-aware DAO techniques and refine the techniques based on our initial results. We will investigate the performance of SMP-aware DAO techniques in the presence of message passing memory traffic and explore techniques that target SOC and PIM in the third year of the project.

Sapphire: Scalable Pattern Recognition for Large-Scale Scientific Data Mining

Chandrika Kamath
Center for Applied Scientific Computing

Summary:

There is a widening gap between our ability to collect data and our ability to analyze it. This problem of data overload has become a serious impediment to scientific advancement in areas as diverse as counter-proliferation, the Advanced Simulation and Computing (ASCI) department, astrophysics, computer security, and climate modeling. To improve the way in which scientists extract information from their data, we are developing a new generation of tools and techniques based on data mining.

Data mining is the semi-automated discovery of patterns, associations, anomalies, and statistically significant structures in data. In the first step of data preprocessing, high-level features are extracted from the data; in the second step of pattern recognition, the features are used to identify and characterize patterns in the data. In this project, we have developed scalable algorithms for the pattern recognition task of classification. We have improved their performance, without sacrificing accuracy. We have demonstrated these techniques using an astronomy application, namely the detection of radio-emitting galaxies with a bent-double morphology in the FIRST survey.

In FY2001, we focused on three tasks: (a) improving the performance of decision tree algorithms, (b) identifying bent-double galaxies in the FIRST survey, and (c) incorporating our research into software to make it easily accessible to LLNL scientists. In decision trees, we focused on ensembles of trees, where the results of several trees are combined through simple voting. We invented two new ways of creating ensembles by randomizing the decision at each node of the tree. The first approach uses a random sample of the instances for each feature. The second approach uses histograms and randomly selects a split point in an interval around the best bin boundary. Using public-domain data sets, we showed that both techniques were more accurate than single trees and competitive in accuracy with other techniques for creating ensembles, but faster.

For the bent-double problem, we focused on galaxies composed of three blobs. Using principal component analysis and exploratory data analysis techniques, we first identified the key features discriminating bent-doubles from non-bent-doubles. This reduced the number of features from 103 to 31. We next input these features to our decision tree software as well as the generalized linear model software from S-PLUS. Varying the number of input features, we created three models for each of the two classifiers. These models were used to classify unseen galaxies and the results were communicated to our collaborators on the FIRST project. Galaxies where all six models agreed were considered bent-doubles with high probability, while those with some disagreement resulted in a lower probability.

We completed the Beta version of our software in December 2000 and Version 1.0.0 in September 2001. This includes the recent algorithms developed for ensembles and the evolutionary algorithm-based oblique decision trees that we had developed last year. We also enhanced the decision trees with several pruning options, splitting criteria, and split finders.

Summary (continued):

Chandrika Kamath
Center for Applied Scientific Computing

We co-edited one book, published 17 papers in conferences, journals, and books, and presented our work at 15 conferences and workshops. We also filed three new records of invention and three patent applications based on earlier work in FY2000. We co-organized two workshops and one week-long program at the NSF Institute for Pure and Applied Mathematics at UCLA, gave three tutorials at conferences, and actively participated in university collaborations.

ISCR

Institute for Scientific Computing Research

Student Internship Research Summaries

CASC Center for Applied Scientific Computing

Out-of-Core KD-Partitioning of Large-Scale Data Sets for Distributed Volume Rendering

Janine Bennett
University of California, Davis

Summary:

There is need for new ways to visualize large-scale data sets. Current methods can be very slow and inefficient since the data sets are larger than can fit in core memory. By dividing up the data sets into smaller load-balanced parcels, and by distributing them over a cluster, we can speed up visualization times dramatically.

Because of memory constraints, an out-of-core kd-partitioning algorithm is used. The vertices of the data set are partitioned spatially into three-dimensional buckets. Further partitioning is then done based on these counts rather than on the actual vertices themselves. If the granularity of the buckets is insufficient to achieve a good balance, the buckets are subdivided recursively until a good balance is achieved. The data in each of the final partitions is stored on disk as an Mpack block. In addition, MPI code was written to transfer these Mpack blocks between nodes on a cluster; this is used to enable load-balancing. Peter Williams, Nelson Max, and Mark Duchaineau were all involved in different stages of the project.

This code will now be used to partition large data sets into smaller blocks which can each be volume rendered in parallel on a node of a cluster.

Parallel Isosurface Refinement

Rita Borgo
University of Pisa, Italy

Summary:

Multi-resolution data-structures and algorithms are key techniques in Scientific Visualization used to achieve real-time interaction with very large data sets. Most research has been primarily focused on the off-line construction of multi-resolution representations, mostly using decimation schemes. Drawbacks of this class of approaches include the inability of maintaining interactivity when the displayed surfaces change frequently. Moreover, it is difficult to guarantee the consistency of the geometric embedding (no self-intersections) of any approximated level of detail of the output surfaces. In our project we have been working on a solution to such problems restricting our attention to the case of meshes computed as isocountours of 3D scalar fields. We adopt a multi-resolution data-structure for the output surface that allows trading accuracy for speed, thereby achieving interactive response times in graphical display. The main feature of our scheme is the ability to extract in a multiresolution fashion the input volumetric data to build a multiresolution version of the output by successive refinements. We use a refinement primitive that updates a local portion of the output so that a consistent output mesh is maintained at any given time. This allows for asynchronous termination of the computation within a small constant delay.

To realize our progressive refinement algorithm and data-structure we use the edge-bisection refinement scheme, which is widely used in the meshing community. This scheme allows one to build a multiresolution data-structure for regular grids without preprocessing since it is equivalent to a predetermined order (octree-like) of reading the vertices. The 3D mesh partitions the region of space of interest into tetrahedra. Each vertex is associated with an input function value. Inside each tetrahedral cell a linear function is used to interpolate the function values at the vertices. In this way we have a piecewise linear representation of the scalar field necessary to compute an isocontour. The coarse level of our refinement is represented by tetrahedra; each refinement step inserts a new vertex on an appropriately chosen edge, splitting each tetrahedron incident on such edge in two halves. The isosurface is updated within each tetrahedron. In this way, as the edge-bisection algorithm makes progress, new function values are introduced and a more detailed definition of the function is obtained. Instead of recomputing portions of the contour, we augment its representation generating directly a progressive data-structure. The refinement can be applied locally to perform adaptive refinement or globally to increase uniformly the resolution of the mesh. This progressive representation of time isocontour can be traversed adaptively independently from the underlying mesh. This kind of approach enables a new kind of tradeoff between speed and accuracy. In our prototype implementation we uncoupled the isosurface construction from its display. The isocontour hierarchy is built by one process that traverses the input 3D mesh. A second process that never accesses the 3D mesh performs the isocontour traversal arid display. For the display part we take advantage of a parallel computation scheme partitiong the data in an efficient tradeoff between load balancing and timing constraints.

We are currently working on the design of alternative traversal strategies of the tree-like structure to improve efficiency of the refinement and to provide a consistent construction of the output mesh in parallel and with minimal interprocess communication. In particular the goal is to design the isosurface construction and a one-way stream of data so that it can work reliably, with no need of synchronization.

Topology-Preserving Subdivision

Peer-Timo Bremer
University of California, Davis

Summary:

In the field of scientific visualization, the computation of isocontours remains one of the most important tools. Of special significance is the topology of isocontours. During visualization no artificial topology should be created and no existing topology should be destroyed. During direct isocontour extraction, these requirements are fulfilled if a sufficiently small numerical error bound can be guaranteed. However, many common data sets become too large for a direct isocontour extraction, especially for real-time applications. The most common solution to this problem is to compute a multiresolution hierarchy in the form of a wavelet transform of the data. The hierarchy can then be used to speed up the isocontour extraction. However, wavelet transforms also introduce artificial topology, especially in the lower resolution representations of the data. For example, new minima and maxima are created much like the over- and under-shoots known from spline approximations. This new topology distorts the real data and can mislead scientists. Our goal is to find a subdivision scheme that allows us to have large differences in resolution while maintaining the correct topology.

To explore the possibilities of topology-preserving and topology-controlled subdivision, we use two-dimensional height fields. As the underlying subdivision scheme we chose the "4-8 subdivision" as described by Velho and Zorin. The 4-8 subdivision is based on semi-regular meshes and bisection refinement. The advantages are the comparatively slow growth factor in each subdivision step and the high-resolution differences possible without creating cracks. However, there has been no method of computing the wavelet coefficients for an adaptively refined 4-8 mesh. Even though the masks that are used during subdivision are small, the evaluation requires a nearly uniformly subdivided mesh. We have developed a new lazy evaluation scheme based on guessing necessary values and correcting possible errors later. Our method allows us to use a fully adaptive 4-8 mesh that uses approximated data at places where data is missing. If the correct data is supplied (a necessary value inserted) the resulting change is distributed over the mesh to correct the difference between the former guess and the real value. If all necessary information is introduced, our mesh is the same as one created by uniform subdivision.

Using the above technique we are developing a subdivision scheme where new topology is inserted only at specific user-defined points. The insertion process is guided by a hierarchy of Morse complexes that describe the topology of the data for different error bounds. The mesh is controlled by the Morse hierarchy in the sense that only the topology of the corresponding Morse graph exists in the mesh. For isocontour extraction this means that no isocontour is lost and no artificial isocontour is created. We avoid creating additional topology by enforcing the following rule: No vertex in the mesh that is not part of the Morse graph (is not an explicitly noted critical point) is allowed to be a critical point. This condition can be checked by visiting the neighborhood of a vertex. However, enforcing the condition without distorting the mesh too much is a challenging problem.

Future research will focus on combining the current adaptive refinement with a full implementation of the topology-preserving algorithms and on methods to control the topological changes. The long-term goal is to lift the scheme to three dimensions.

A New Parallel Coarsening Method for Algebraic Multigrid

Oliver Broeker
Swiss Federal Institute of Technology

Summary:

Algebraic multigrid (AMG) is known to solve a large class of linear equations arising from the discretization of partial differential equations. Part of the algorithm is a greedy procedure to choose the coarse unknowns for a given system. This procedure is inherently sequential. Parallel implementations of an AMG typically use an approach where sequential coarse grid choice is done on the local grids and boundary points are treated separately. The resulting convergence rate is dependent on the number of processors. We seek a parallel coarsening process that is free of this limitation.

We devised an algorithm based on ideas similar to the "compatible relaxation" of Brandt. A Matlab implementation yields good convergence rates and bounded operator complexities for a variety of small anisotropic diffusion and convection diffusion problems. The algorithm was added to the BoomerAMG code and tested on larger matrices.

The method needs further investigation, especially since the choice of its threshold parameters is crucial and nontrivial. The work will be continued at the ETH as the focus of a doctoral dissertation.

Using Meta-Data to Automatically Wrap Bioinformatics Sources

David Buttler
Georgia Tech

Summary:

A large amount of bioinformatics data is distributed over the Internet. Typically, this information is accessible only through custom, web-based query interfaces. These interfaces often include features uncommon in industrial applications such as multiple parameters, a variety of query options that invoke different support programs, indirection and delay pages, and complex results. If users require information from multiple sources, they must pose the appropriate queries at each source individually then explicitly integrate the results. This solution may be acceptable for a small number of sources, but it quickly becomes an overwhelming burden for users. Currently there are over 500 bioinformatics sources for biologists to choose from, making it infeasible to manually gather data from a non-trivial fraction of the available sources.

Our goal is to simplify access to bioinformatics data by providing a single access point to a large number of sources. The fundamental problem is that each source is fiercely independent. As a result, they use different semantics, customized interfaces, and unique data formats. Furthermore, they are prone to having their interfaces and formats frequently updated without warning. In order to maintain access to a source, a specialized access program (wrapper) that keeps pace with the source's evolution is required. We present a general meta-data format capable of describing the complex, web-based interfaces often found in bioinformatics. Because this description provides sufficient information to automatically generate a wrapper for the associated site, it is an important first step in developing an infrastructure capable of interacting with a large number of dynamic, heterogeneous data sources.

Rather than create a new meta-data format from scratch, we started with the DAML specification. DAML is an extension of XML and RDF that allows the definition of machine understandable ontologies. Specifically, we have extended DAML's web service description language, DAML-S. A service in DAML-S is essentially a web-based interface to data. The DAML-S description of a service has three parts: a service presents a profile, which indicates what information the service requires (input) and what the service does (output); it is described by a service model, which indicates how each step of the service works; and it supports a service grounding, which indicates how to access the service. For the purposes of our data source description, a profile is the external interface to a web service, the model represents the components of the service, and the grounding is the detailed information about how to properly interact with the service interface.

The DAML-S model describes how components of a service interact. The most important characteristic of this description is the Process, which has inputs, outputs, and effects (real world consequences for invocation). Processes can represent atomic actions, or control flows (e.g., looped, conditional, sequenced [ordered], and parallel [unordered] execution). To complete the interface descriptions of data sources, we have extended DAML-S to include a detailed specification for groundings (the current release does not define them). Our definition of a grounding contains several elements: (1) a pointer to the Process it grounds; (2) the binding; (3) a specification of how to convert Process inputs into parameters the binding can use; (4) a parser; and (5) links.

Summary (continued):

David Buttler
Georgia Tech

We hope that this work will speed the evolution of the DAML-S standard and help establish a common format for complex interface descriptions. However, our primary motivation for this work has been to make the first steps towards an infrastructure that can support access to a large number of dynamic, heterogeneous scientific data sources. As such, we are continuing to develop this infrastructure by extending the XWrap Elite wrapper generator to use our metadata descriptions, instead of human interaction, to generate a wrapper for the associated source. Once completed this program will generate full-featured wrappers that conform to a common query interface and return objects in XML. Providing this uniform source interface will enable different integration scenarios, from simple multi-source result fusion to techniques that reconcile semantic information such as data warehouses or federated systems, to be explored. Thus, it is an important first step in providing a single access point for bioinformatics researchers.

A Generic Scheduling Simulator for High-Performance Parallel Computers

Gyu Sang Choi
Penn State University

Summary:

Job scheduling is very important to meet versatile user demands in large-scale, high-performance computing environments. Since a systems environment is constantly changing, redesign or reevaluation of scheduling policies is frequently needed. Generally, system designers choose the simulation approach to evaluate different scheduling policies. The development of a new scheduling simulator, however, is a time-consuming process. Nevertheless, the simulator cannot be reused easily for a new environment. Thus, it is very crucial to develop a scheduling simulator independent of environments.

We propose a generic scheduling simulator for high-performance parallel computers. The simulator is based on an M-square framework for representing various systems. We expand the M-square framework because it has some limitations for a generic scheduling simulator. The proposed simulator supports various homogeneous machine architectures, takes various workloads, makes a user describe easily a new scheduling policy, and calculates memory and context switching overhead. Therefore, a system designer can avoid the overhead to develop a simulator, using the simulator.

There are some limitations in a current general scheduling simulator. This simulator does not model dynamic events like message passing and synchronization mechanisms, mainly due to the lack of a mechanism to represent inter-task communication patterns. We will expand this simulator to support dynamic events in order to run the simulation for implicit co-scheduling or dynamic co-scheduling. Currently, the simulator supports only a homogeneous environment, but we plan to support heterogeneous environments in the next version. The generic simulator can measure a memory overhead but does not run the simulation with a network model. Since this issue is related to message passing and synchronization mechanism, these two issues will be manipulated together in the future.

Two-Grid Methods for Radiative Transfer

Todd Coffey
North Carolina State University

Summary:

Nonlinear radiation diffusion problems can be notoriously difficult to solve implicitly. Application of nonlinear solvers requires repeated solution of a linearized version of the problem, and these linear problems are often the bottleneck for the entire solution process. We are trying to improve on these traditional methods by developing a two-grid scheme in which the problem is solved very accurately on a coarse mesh followed by a single solve of a linearized version on the original problem mesh. The motivation for the method is that the coarse-mesh solution will capture the nonlinear behavior of the problem. Thus, when the solution is moved to the fine-mesh we need only refine it with a linear solve. As a result the nonlinearities are dealt with on a much smaller mesh than the original fine mesh, resulting in a faster solution process. This method was successfully applied to a similar problem arising in ground water applications by Wu and Allen in 1999. The main theoretical development for this method comes from J. Xu (*SIAM Journal of Numerical Analysis*, 1996) and the extensions to our nonlinear system come from C.N. Dawson, M.F. Wheeler, and C.S. Woodward (*SIAM Journal of Numerical Analysis*, 1998).

We implemented this method by modifying an existing code for solving this problem. The existing code employs LLNL's PVODE as an adaptive time-stepping ODE solver (BDF time integrator and Newton nonlinear solver), transport3d (a radiative transfer in-house development code) providing the right-hand side function along with *hypre* (LLNL's high-performance preconditioner library) helping with parallelization and the preconditioner. We modified PVODE to handle two grids and added that capability to transport3d. With these changes, we could then solve the coarse mesh problem using the same machinery as before. The grid transfer routine was then written using trilinear interpolation on the transport3d side. Finally, a fine-grid linear solve will be written that will use the preconditioned GMRES solver from PVODE and a new right-hand side function from transport3d. The changes to both codes abstract the method so that it can be applied to different problems. Eventually this method may migrate to KINSOL and IDA (LLNL's nonlinear system solver and an implicit DAE system solver).

We expect this two-grid method to be especially effective in very nonlinear problems where the nonlinear solver in PVODE requires many iterations. There are several physical parameters in the radiative transfer problem that we can adjust to make the problem more nonlinear. We will conduct a series of experiments to determine how this method compares with the existing solution method.

Computations in Scalar Field Topology

Kree Cole-McLaughlin
University of Utah

Summary:

Scalar fields are used to represent data in different application areas like geographic information systems, medical imaging, or scientific visualization. One fundamental visualization technique for scalar fields is the display of isocontours, that is, sets of points of equal scalar value, also called the isovalue. For example, in terrain models isolines are used to highlight regions of equal elevation. The Contour Tree is a graph that represents the relations between the connected components of the isocontours in a scalar field. Two connected components that merge together (as one changes continuously the isovalue) are represented as two arcs that join in a node of the graph. Each node occurs at a critical point in the scalar field. For a simphicial domain with a piecewise linear interpolant, Carr et al. present a simple algorithm for computing the Contour Tree with complexity O(nlogn+ N), where n is the number of points in the domain, and N is the number of simplicies.

One fundamental limitation of the Contour Tree is the lack of additional information regarding the topology of the contours. The topology of an isocontour is fully characterized by its Betti numbers. In 3D fields the isocontours are surfaces and the Betti numbers correspond to the number of connected components, the number of tunnels, and the number of voids enclosed by a surface. Pascucci has presented a simple and fast algorithm for the construction (if the Contour Tree of a 3D scalar field augmented with the Betti numbers of each contour). The complexity of this approach is O(N log N) in time and O(N) in storage. The nodes of the Augmented Contour Tree correspond precisely to every critical point of the field. Thus, we trade a slightly slower algorithm for more information.

We have implemented both of the above algorithms and produced images of complex contour trees for a few datasets. The images are produced with the graphviz utility from AT&T. Our datasets are quickly approaching the limit of this utility's ability. Therefore, in the future we will need to develop our own tool for drawing the Contour Tree. We have designed and implemented a third algorithm, which has not been published yet, that computes the Augmented Contour Tree in O(n log r+N) time. This new algorithm combines the efficiency and simplicity of the Carr scheme in with the augmented information introduced in the Pascucci scheme. Thus we have successfully introduced new and useful information to the Contour Tree data structure without slowing down the computation.

We have several goals for future development of this project. Currently we are working on scaling the computation up to very large datasets by working out an algorithm that can be implemented on parallel computers. This new approach has an added advantage in that it does not require a simplicial domain or a piecewise linear interpolant. Hence it is applicable to the trilinear interpolant that is commonly used for rectilinear grids. We are working on improving the complexity of the current algorithm by eliminating the O(n log n) term. There are plans to implement the Contour Tree as an interface to visualization software. This opens up a whole range of possible information that can be presented for each contour. We are also working on the problem of drawing the Contour Tree for time-varying fields.

Parallelization of the HIAC Volume Renderer and Efficient Determination of an Exact Visibility Ordering

Richard Cook
University of California, Davis

Summary:

The HIAC volume renderer is software that does volume rendering of unstructured grids. The zoo cells formed by an unstructured grid are first sorted into a visibility order and then projected onto the viewplane using an algorithm by Nelson Max and displayed using OpenGL hardware rendering.

The projection step is the slow step in the program. Therefore, in this project the projection and rendering steps were to be parallelized to improve performance on large datasets.

The parallelization of the projection step was done using pthreads under a shared memory processor architecture. The work process is pipelined into sort, projection, and OpenGL steps, with multiple projecting threads, as that step is the bottleneck. Each projecting thread is added to a queue of consumers and one or more threads at the head of the queue are awakened each time the producer finishes a quantum of work. As each finishes, a final rendering thread converts the information stored by the projection threads directly into optimized OpenGL hardware calls. Near linear speedup was obtained for limited numbers of threads, but scalability in the sense of constant speed for increasing data set sizes has not been demonstrated.

During the parallelization effort, we also created and implemented a replacement for the MPVO sort by Peter Williams. The new Scanning Exact MPVO (SXMPVO) algorithm performs an exact (to the resolution of the resulting image) visibility ordering of the zoo cells and is very fast.

The parallelization and visibility ordering work was significant enough to result in a Master's thesis for Richard Cook and two paper submissions to the *IEEE Visualization Conference* in 2001.

HIAC can be integrated into a user tool that is scalable and robust to be used by scientists at the Laboratory to generate high-accuracy images of their data. This may require re-engineering HIAC's research code using a more modular and extensible approach.

Overlapping Finite Elements for 2D Linear Elastic Problems

Nathan Crane

University of Illinois, Urbana-Champaign

Summary:

Standard finite element analysis is performed on a fully compatible mesh. This implies that inside a continuum body all nodes occur only at the corners of an element and that there are no gaps or overlaps between elements. CAD packages generally do not construct a compatible model. Instead a CAD package produces a solid from the union, intersection, or subtraction of simple parts. Each of these simple parts such as cylinders, NURBS surfaces, and trimming curves are individually very easy to mesh. However, producing a compatible mesh from these intersecting simple parts is very difficult and often requires extensive human intervention. An alternative mesh construction strategy is to not bother producing a compatible mesh but instead solve the problem on a set of overlapping meshes.

We produced a scheme that creates a valid 2D mesh from overlapping elements. This method is implemented by creation and constraint of boundary nodes. This scheme is compatible within the framework of a standard iterative solver. Correctly formulated overlapping meshes will pass the "patch test." Thus, with sufficient refinement the exact solution to a problem can be obtained. Experiments on complex geometries demonstrate that overlapping meshes are as accurate as a corresponding compatible quad mesh and much superior to a triangle element-only mesh. Overlapping meshes do require more elements and more degrees of freedom than a corresponding compatible mesh. Though a overlapping mesh may be computationally slightly less efficient, it can be constructed with less human interaction, reducing total problem solution time and effort.

Parallel Gauss–Seidel Methods

Paul Dostert
Texas A&M University

Summary:

We are interested in numerical solutions to the general sparse matrix problem, $Ax=b$. Typically, A comes from some finite element discretization of an elliptic PDE. A Gauss–Seidel algorithm is frequently used as a smoother or preconditioner for many numerical techniques which solve $Ax=b$. Often, general solution methods are applied on extremely large sparse matrices, making parallel algorithms a necessity. Unfortunately, the Gauss–Seidel algorithm is not easily parallelized because each point in a grid needs recently updated values of some previous points. We investigate different distributed memory parallelizations of a Gauss–Seidel algorithm for unstructured grids.

In order to begin any type of parallelization for Gauss–Seidel, we must first use a coloring technique. The idea of coloring is that we may run different pieces of a grid that never communicate directly with one another, simultaneously. Two different coloring techniques were used for our algorithm: nodal coloring and domain coloring. Nodal coloring involves assigning each node in our domain a color such that no two neighboring nodes have identical colors. We then run through the colors sequentially, but may distribute the nodes of any given color over many processors. One problem is that since the colors are spread out throughout our domain, passing these individual nodes between processors is a time-consuming task. This prompted us to look at domain coloring.

To perform domain coloring we must first decompose our grid into multiple non-overlapping domains. This is accomplished using MeTiS, an element-based splitting program. Differences between two MeTiS versions, pmetis and kmetis, were also analyzed. Once we split our domain, we use the same technique as with the nodal coloring, only we color whole domains now. Therefore, we color the domains such that no two neighboring domains have the same color. One large advantage over nodal coloring is that only the elements that are in multiple domains must be passed.

All of our techniques were programmed in C using MPI. The algorithms, for large problems, were run on the Compass Clusters and ASCI Blue Pacific. The reduction of passing from nodal to domain coloring increased the speed of the Gauss–Seidel almost ten-fold for some domains. We were also able to obtain speed increases as we used more processors for the domain coloring code.

Possibilities for further work include switching from a traditional Gauss–Seidel algorithm to one with overlap (Schwarz). This will involve decomposing our grid using MeTiS, then creating an overlap between the domains. Once this overlap is made, the new domains with overlap will be colored, and then run in parallel. A serial overlap algorithm converges in approximately half the iterations of a traditional serial Gauss–Seidel. Our hope is that the increase in passing and work done per iteration will be offset by an increase in convergence rate, therefore decreasing overall computation time.

Optimal Shape Design for a Layered Periodic Helmholtz Problem

Mike Flanagan
Texas A&M University

Summary:

A standard model of the Helmholtz equation was studied, using non-regular elements applied to an h-p refinement mesh. One of the problems inherent in the discretization of the Helmholtz equation is unwanted reflections in the outer boundary. An attempt is made to resolve the function to its natural analytic state as we move towards the outer boundary. The idea was to use a simple geometry for the domain and develop a triangulation (here, rectangles were used with tensor product basis functions) scheme that would allow for the p-refinement to increase in order as we move further from an internal boundary, being a square in this case. Also, the mesh size increases as we move out, with the essential rule being that two elements' faces form a side of the next face as we move out in layers from the internal boundary. The boundary conditions employed at the far end are standard Sommerfeld Radiation conditions. We have developed a two different mesh schemes, and further work will include studying the associated mass-matrices resulting from such a scheme. Issues to be resolved are potential numerical instabilites associated from the larger degree polynomials.

Low Machnumber Compressible Flow Simulations Using a Multigrid Procedure

Achim Gordner
University of Heidelberg

Summary:

Compressible flow simulations in a low Mach number regime suffer under the existence of acoustic waves. In particular, long wavelength pressure fluctuations are coupled with short wavelength velocity fluctuations. Since it is our intention to perform aeroacoustic flow simulations, the accurate resolution of the above-mentioned coupling is essential for an appropriate simulation. Having this multiscale coupling problem in mind, a multigrid solution procedure is chosen. However, standard multigrid procedures cannot be applied, since the resulting system of algebraic equations tends to become stiff if the Mach number goes to zero. We concentrate on modifications introduced into the smoother to make it applicable to resolution of the pressure-velocity coupling.

Recent approaches include using the multiple pressure variable technique. Since therein a certain decoupling between the pressure and velocity fluctuations is mandatory, we decided to use a full approach using the entire compressible Navier–Stokes equations for an ideal gas. We extended the primitive variable discretization based on our experience for incompressible flow simulations to the compressible case. Since off-diagonal matrix entries are dominant and, therefore, should be taken into account within the smoothing process, a strong incomplete LU factorization was chosen to act as smoother within the multigrid cycle. Convergence of the smoother was reached for low Mach numbers with appropriate numbering of unknowns and introduction of some extra fill-in. However, developing an optimal algebraic scheme limiting necessary fill-in is still under way. To prove that the implementation of the incomplete LU factorization is still applicable in parallel computations for this application, we employed the CASC Compaq cluster.

Modifing the smoother within the multigrid cycle to count for the multiscale coupling is one possibility for adapting the multigrid solution procedure. Another solution may be the proper choice of restriction and prolongation operators.

Interactive Exploration of Large Iso-Surfaces in Volume Datasets

Benjamin Gregorski
University of California, Davis

Summary:

Numerical simulations performed on LLNL supercomputers are generating unprecedented large amounts of data. A standard method for visualizing, exploring, and understanding this data is to examine various isocontours. The problem with directly viewing the isocontours is that they contain hundreds of millions of triangles. Surfaces of this size cannot be viewed at interactive frame rates on conventional desktop machines. Intelligent preprocessing techniques and runtime algorithms are needed for interactive exploration and visualization. The goal of this project is to develop a system for dataset exploration on desktop workstations. Large datasets are divided into subsets called bricks and stored on disk or on a remote system. These bricks are loaded and unloaded by the application as they are needed so that system resources are effectively utilized.

Interactive frame rates can be achieved using algorithms that employ multiresolution data structures with view-dependent rendering. Multiresolution data structures make better use of available storage space and computation power by representing important areas with more detail and less important areas with less information. View-dependent rendering selects what portion of the data to render based on where the user is looking and on the user's perception of the data. Objects outside the field of view do not need to be rendered, and objects that are far away can be rendered at lower resolutions. View-dependent rendering takes advantage of the multiresolution data structure to efficiently select what to draw.

Our multiresolution data structure is a recursive tetrahedral mesh based on longest edge bisection. The mesh structures supports fast, local refinement necessary for view-dependent rendering. In addition, it supports an efficient method for ensuring mesh continuity required for iso-surface extraction. View dependent rendering calculates the distortion on the view screen; it adjusts the refinement of the multiresolution mesh, selecting finer levels where more detail is needed and coarser levels where less detail is needed.

The basics of the multiresolution data structure and view-dependent refinement have been implemented. Our continuing work is focused on working with large datasets, accurate representation of the data, and computational optimizations.

The Fast Multipole Method for Computational Electromagnetics

Chaz Hales
Brigham Young University

Summary:

Integral equation methods, such as the Method of Moments, are useful in computational electromagnetics because they accurately model the electric or magnetic field resulting from a given source distribution. Because the field resulting from a point source is well known, we can easily find the field from a source of arbitrary distribution by integrating this "point source" expression over the desired source distribution. The resulting integral equation is in fact a linear system of equations and can be solved in a straightforward manner. An important drawback in solving these linear systems of equations is the fact that the matrix describing the interactions between these source points is a large, full matrix. The time required to find the solution for such a system is on the order of *N*-cubed for direct solution methods, where *N* is the number of points, or unknowns. Even faster iterative solution methods require *N*-squared computations for each iteration. As we desire to solve larger and larger problems (that is, as the number of unknowns *N* increases), we are hindered by this computational bottleneck.

The Fast Multipole Method (FMM) yields computational savings by decreasing the number of computations required per iteration of the solution method. At the core of FMM is the idea that direct interactions between one point and every other point can be adequately accounted for by first grouping the points into smaller clusters. Rather than computing the interactions between all the points directly, only the interactions between these groups are computed, thus reducing the number of computations required to solve the system.

Currently we have implemented FMM for a very simple electromagnetic problem. We have used FMM to solve the linear system resulting from a perfectly conducting (PEC) sphere illuminated by a simple planewave in free space. This canonical problem serves as a benchmark that allows us to compare initial FMM results to an exact analytical solution. Accuracy of three decimal places was achieved in test cases, and it is hoped that additional accuracy can be achieved in future test runs.

Immediate plans for the future include extending this simple PEC case to allow us to model other materials (dielectrics). Also a Multi-Level Fast Multipole Algorithm (MLFMA) will be implemented allowing us to achieve even greater computational savings. The net effect of these changes will allow us to solve a broader category of problems, as well as solve larger problems than currently can be solved in reasonable run time.

Adaptive Mesh Refinement for Oceanic Carbon Sequestration Simulations

Aaron Herrnstein
University of California, Davis

Summary:

The continuously increasing rate of CO_2 emissions into the atmosphere is an ever-growing concern of climatologists. It is hypothesized that an excess of such gases will produce a "greenhouse" effect, thus increasing global temperatures and ultimately altering global climate. As an alternative to releasing greenhouse gases into the atmosphere, it has been proposed to deposit them in the ocean. Major concerns of this carbon sequestration are the amount of CO_2 that escapes back into the atmosphere and biological impacts brought about by possible changes in ocean pH.

Ocean General Circulation Models (OGDM) are used to model sequestered carbon over a time scale of centuries. The OGCM used by LLNL's Climate and Carbon Cycle Modeling group shows discrepancies in pH changes between fine and coarse grids. A finer grid is needed to determine convergence, but refining the entire grid will be quite costly in terms of computer time (on the order of months in running time). Adaptive Mesh Refinement (AMR) allows desired sections of the grid to be refined, thus reducing run time considerably.

It is the goal of this research to produce an OGCM that uses AMR. Such a tool will be very useful for areas such as Carbon Sequestration. To aid AMR implementation, the software package SAMRAI (Structured Adaptive Mesh Refinement Application Infrastructure) is used. Refinement of the tracer field is based on gradients and possibly other criteria as needed. In addition, regions of critical topography might be refined if necessary. Current capabilities include a fully functional tracer advection model using AMR and a diagnostic meshing code that uses data from a traditional OGCM. Work for the immediate future involves creation of a shallow water dynamics model using AMR.

Euclid Parallel Preconditioning Library

David Hysom
Old Dominion University

Summary:

Euclid is a scalable parallel ILU preconditioning library that supports arbitrary levels of fill. The implementation is based on our PILU algorithm that was presented at SuperComputing 1999, and later published in expanded form in the *SIAM Journal on Scientific Computing*. This summer brought the work to fruition, with many experimental studies conducted on the ASCI Blue platform, and the code's inclusion in the LLNL *hypre* library.

The algorithm takes a matrix as input and assumes no *a priori* knowledge of structure, symmetry, or other unique features. The only requirement for good performance is that the graph of the matrix be partitionable such that each subdomain has a large interior to boundary node ratio. This requirement is met in practice by most mesh and problem generation codes in the physical sciences.

This summer's experiments centered around two key questions. First, how does the preconditioning technique scale when problem size is increased in proportion to the available processing power; second, how does PILU performance compare to Block Jacobi preconditioning (Block Jacobi being a well-known preconditioning technique that is trivially parallelizable but not scalable in convergence rate). The following representative experimental results are for the IBM-SP Blue Pacific machine at LLNL.

The triangular solve (a.k.a., preconditioner application) phase scales on a per iteration basis, as problem size is scaled in proportion to the number of processors. In one of our examples, a 2D convection-diffusion problem, each processor owned a grid of approximately 65 thousand unknowns. Global problem size was increased from 65 thousand unknowns (one processor) to 26 million unknowns (400 processors). We tested factorizations with various fill levels, e.g., ILU(1), ILU(3), and ILU(6). In all three cases the lines "ramp up" in the range of 1 to small numbers of processors. This is expected and is attributable to "filling up" the communication pipeline. The near flatness of the execution time performance for increasing problem size from 64 through 400 processors for all levels indicates near perfect scalability.

Tests on a simplified 2D nonlinear radiative transport problem on up to approximately 4 million unknowns on 400 processors compare our new PILU preconditioning technique with the well-known and highly parallel Block Jacobi technique. Results show that PILU preconditioning is increasingly advantageous as problem size increases. For 400 processors, PILU preconditioning cut total solution time by 50%. PILU is not convergence-rate scalable as is, for instance, multigrid. However, it is a very strong preconditioner with excellent per-iteration scaling properties. Hence, it could be embedded in a multilevel algorithm as a smoother.

Using Genetic Algorithms for Image Processing

Ty Jones
University of Nevada, Reno

Summary:

The Sapphire group is working toward an automated image understanding system. Image processing is a very important part of image understanding. To automate such a system, the use of genetic algorithms is potentially very useful. We decided to try implementing the Phoenix segmentation algorithm with genetic algorithms. This approach was attempted first by Bir Bhanu with promising results.

The first step was to implement the Phoenix algorithm. The algorithm utilizes the histogram of the image to divide the image into smaller regions. A new histogram for each new region is used to subdivide further. This is a recursive process and the algorithm requires input of a large number (17) of individual parameters. These parameters are what we wanted to optimize with a genetic algorithm. The implementation involved the use of much of the software previously developed in Sapphire including an evolutionary algorithms package and various other data structure templates. Although the system was not completed, we are expecting to get significantly better results than if we were to just use some default input parameters. Most of the implementation details were decided upon in conjunction with Chandrika Kamath. The evolutionary algorithm package was developed by David Littau, and much of the data-structure template structure was developed by Nu Ai Tang and Erick Cantu-Paz.

Building a Cache Component for the Memory Wall Project

Michael King
University of Utah

Summary:

Memory speed is growing at a much slower rate than processor speed. As time passes, the difference between the two speeds increases, making good cache design increasingly important. Because of this gap in speed, programs running on a multiprocessor system do not always make use of the high-processing rate potential. Our goal is to explore ways to optimize the use of cache memory in a multiprocessor environment.

Our approach is to build a trace-driven simulation of a memory system under representative applications in order to find out what optimizations to try on a real machine. Since we are not certain of every detail of the system we are simulating, it must be reconfigurable. We need to be able to change the hardware components that go into the simulation, as well as know how each component links to the others.

While in residence, I have built a reconfigurable cache component that simulates a generic cache. The simulation takes memory addresses as input and outputs statistics such as the hit/miss ratio and timing information. This tool should be useful in evaluating reconfigurable cache designs in the ongoing LDRD Memory Wall Project.

Sensitivity Analysis Techniques for Radiative Transfer

Rachel Knop
West Point Academy

Summary:

Radiation diffusion problems can often be modeled by a system of parameter-dependent differential equations. In many cases, it is not only important to solve the differential equations but also to quantify the sensitivity of the results to the problem parameters. These sensitivities can then be used to determine which parameters are most influential in affecting the simulation results.

The goal of this project was to compare methods for computing the sensitivity of a radiation diffusion problem with respect to a specific parameter. The model was comprised of a coupled set of differential equations for radiation energy and material energy. The model was numerically solved using PVODE, an ordinary differential equation solver. The parameter we considered was actually a scalar value used to multiply the Planck opacity values that appear in the coupled equations. The nominal value for this multiplier was 1.0. The so-called "brute force" method computed Planck sensitivity by taking the difference between the nominal results and results obtained when the multiplier was perturbed slightly, and then dividing these differences by the value of the perturbation. The magnitude of the perturbations ranged from 0.1 down to 0.000001. These sensitivity results were then compared with those obtained using SensPVODE, a sensitivity analysis version of PVODE. The SensPVODE approach involved deriving a differential equation for the Planck sensitivity and then simultaneously solving the radiation energy, material energy, and Planck sensitivity differential equations. With full error control, SensPVODE made step size and order selections based on the behavior of this entire system of differential equations. With partial error control, SensPVODE excluded the Planck sensitivities from the step size and order decisions.

We compared the brute force sensitivity method with both versions of SensPVODE (full error control, partial error control). The three sensitivity methods computed the Planck sensitivity at the same set of output times. Differences between Planck sensitivities were measured in terms of maximum relative differences, with the relevant SensPVODE results used as the reference solution. Initial results indicate that the differences were comparable when the brute force method was compared to the full or partial SensPVODE method. These sensitivity differences were relatively small initially (10% at most), and then again at later times when the radiation energy and material energy approached steady state. The most significant differences occurred at certain output times, roughly one-quarter to one-half of the way through the simulation. The brute force method is appealing because it computes sensitivities at far less cost than either SensPVODE method. The main objective in continuing this study will be to determine whether perturbation values that give brute force sensitivities as accurately as either of the SensPVODE methods can be found.

Parallel Implementation of an Algebraic Mortar Method

Tzanio Kolev
Texas A&M University

Summary:

We developed a parallel code implementing the mortar method with algebraically constructed multiplier spaces. The target application of this code was the construction of parallel multilevel preconditioners using element based (AMGe) coarsening in each subdomain. The work was carried out in close collaboration with Professor Joseph Pasciak from Texas A&M University and Dr. Panayot Vassilevski from CASC at LLNL. Parallel methods for numerical solution of elliptic partial differential equations typically require decomposition of the original computational domain into subdomains that are assigned to individual processors. Generally, the subdomain triangulations (meshes) may not align across the interfaces. The mortar method is a technique that allows for non-matching discretizations by imposing continuity of the finite element solution across the interfaces in a weak sense. More specifically, the jump of the discrete solution on each interface is orthogonal to a mortar multiplier space.

In this project the algebraic mortar technique (a pure algebraic extension of the finite element mortar method) was used to handle the case of parallel element coarsening with application to AMGe. The element based algebraic multigrid (AMGe) is an algorithm to coarsen generally unstructured finite element meshes where only the fine-grid element stiffness matrices are given. The AMGe operates on and produces generalized "elements" defined in terms of relation tables between nodes, faces, and elements. Even if the initial mesh is matching, independent (parallel) AMGe coarsening in each subdomain will produce non-matching meshes. These non-matching meshes can be "glued" by a mortar space to form a global problem. We considered an algebraic extension of the local construction for the mortar multipliers based on the general 3D dual finite element basis described by Kim, Lazarov, Pasciak, and Vassilevski.

A general code was implemented to illustrate the behavior of the proposed method. It requires input data for each subdomain that includes the element topology, the local subdomain stiffness matrices, as well as the mass matrices on the interfaces. This information is independent of the dimension and structure of the problem and is regenerated after an AMGe coarsening.

A number of tests for 3D problems were performed. An initial, generally unstructured tetrahedral mesh created by the mesh generator NetGen was pre-refined, then partitioned with METIS and then refined (now locally) independently in each subdomain to produce a generally non-matching mesh for the whole domain. The refinement was performed by ParaGrid, a parallel refinement code available in LLNL. A visualization tool based on OpenGL that allows for three-dimensional exploration of the solution was also developed.

The implementation constructs the initial mesh and calls an external AMGe coarsening subroutine developed by Dr. Panayot Vassilevski. The communication on the interfaces was implemented in an abstract way that can run on serial machines as well as in parallel based on MPI. The main target application of the code, a parallel multigrid preconditioner for the mortar system on the finest level, is being tested.

The Generation of Optimizing Preprocessors Using the ROSE Infrastructure

Markus Kowarschik
University of Erlangen-Nuernberg, Germany

Summary:

Large application codes in scientific computing are usually based on the use of libraries providing various data structures and functions. This is particularly true for object-oriented applications where these libraries implement high-level abstractions like array classes, grid classes, etc. Unfortunately, the use of high-level abstractions, which are defined in underlying libraries, cannot be optimized by the compiler since the semantics of these abstractions are user-defined and, as a consequence, unknown to the compiler. This lack of knowledge is a major reason for the poor efficiency of many object-oriented scientific codes: the Mflop/s rates that can be measured at runtime are just tiny fractions of the theoretically available peak performances that the manufacturers of the machines claim for their products.

ROSE is a software infrastructure for generating library-specific preprocessors that perform source-to-source transformations, e.g., eliminating the need for creating temporary objects and introducing cache-based transformations into computations on structured grids. Internally ROSE parses the application code (currently C++ code) and assembles the corresponding abstract syntax tree (AST). The use of library-specific high-level abstractions is recognized using high-level grammars, which are derived from the grammar of the base language C++. Since the application might be based on several libraries there might be a whole hierarchy of ASTs. After the ASTs have been modified according to the transformations specified by the library writer(s), the base language AST is unparsed. This final step yields the optimized C++ source code.

The work during this summer has mainly focused on two issues. First, we have investigated techniques for the specification of the source-to-source transformations. We have implemented a specification mechanism that is based on the assembly of optimized C++ code using strings of source code. This source code is then passed to the C++ compiler front-end, which generates the AST fragment to be plugged into the application's AST replacing the original AST fragment to be optimized. Second, we have implemented the automatic generation of a general-purpose tree traversal mechanism, which can be used in various places within ROSE, e.g., in order to simplify querying the AST during the transformation phase and in the course of unparsing the C++ AST after all transformations have been applied.

Parallel Unstructured Adaptive Multigrid for Instationary Problems

Stefan Lang
University of Heidelberg

Summary:

High-level parallel applications based on the message-passing paradigm are difficult to design and implement. This is especially true when solution adaptive techniques are used and problems on complex geometries are faced. Many of these difficulties are addressed inside the UG (Unstructured Grids) platform, e.g., dynamic load migration and load balancing, parallel grid adaption, and parallel I/0 and graphics. The driving idea during the design process of UG, a multigrid code for the computation of partial differential equations, was to find proper abstractions for each of the different functional parts of a parallel, adaptive, and unstructured software framework. This assures a maximal degree of code reuse and treatment of various partial differential equations becomes possible without superfluous coding effort.

Here we describe the usage of the UG library to compute density-driven flow through a ground layering system consisting of 63 different geological layers with permeability values changing by a factor of 10**4.

Time discretization is done with a BDF method. Space discretization is done by an implicit finite volume scheme. Using Newton linearization the nonlinear system is transformed into large sparse linear systems. These are then solved by parallel multigrid with BiCGSTAB. For visualization purposes, a scalar output, the solution in terms of the salt mass fraction is drawn using the parallel graphics device based on distributed z-buffers.

Using 1, 4, 16, 64, 256 and 512 processors in a scaled-sized computation, we computed up to 60 million unknowns. The linear solution process is analyzed for numerical and parallel speedup (efficiency). We show an average increase in number of iterations from 35.5 to 51.8, leading to a numerical efficiency of 0.61. Iteration time increases from 4.2 on 1 processor to 6.9 seconds, leading to a parallel efficiency of also 0.61. Thus, the overall efficiency is 0.61*0.61 = 0.37 and the corresponding speedup is 190. Timings for visualization range from 3.6 on 1 processor up to 4.6 seconds on 512 processors for creating one frame for the final movie.

In the future we will undertake parallel 3D calculations on processor numbers exceeding 10**3 for a variety of timesteps. We will perform scalability analysis of the linear and nonlinear solution process. We expect to be able to analyze up to 10**8 unknowns, a commonly cited ASCI goal, examining especially grid convergence behavior at sensitive points of the computational domain.

Data Classification Using Decision Trees

David W. Littau
University of Minnesota

Summary:

Decision trees are used to classify data. They use a set of labeled data to train, or build, the decision tree, based on the values of the data attributes. There were two kinds of decision trees typically considered: Axis-Parallel (AP) and oblique. An AP tree splits the data based on one numeric attribute, and an oblique tree splits the data using a linear combination of all the numeric attributes. A numeric split results in two new branches. If a given attribute is non-numeric, it is split into one branch for each separate nominal value. The tree construction uses a greedy algorithm that minimizes the local error for each split. Once the tree is built, it can be used to classify all the unlabeled data associated with the labeled data.

The Sapphire library already contained a fully operational implementation of the AP decision tree code. The capabilities of the library were expanded by adding code that created approximate AP trees, oblique decision trees using the OC1 algorithm, an approximate OC1 implementation, and oblique decision trees using an Evolutionary Algorithm (EA). The approximate AP trees differ from standard AP trees in that the data are not sorted to build an approximate AP tree. Instead, a histogram is created and the histogram boundaries are used to calculate split points. The OC1 code implements a standard oblique decision tree algorithm that uses an AP split as a starting point. The approximate OC1 code uses an approximate AP split as a starting point, as does the EA code. The oblique EA code also required some modifications in and enhancements to the existing EA library code. Any desired EA, be it a Genetic Algorithm, an Evolution Strategy, Evolutionary Programming, or a hybrid of all three types, can be used to evolve the oblique splits.

Experiments were performed using the approximate AP trees. An enhancement was added that randomized the split point for each branching, and an ensemble of 10 of these randomized trees was trained. The ensemble classified the data using a majority vote. The results for a variety of data sets indicated that the ensemble was, in most cases, better at classifying the data than a single regular AP tree. More experiments are planned by the Sapphire group, which will result in at least one publication. The new decision tree code will also be used as part of the ongoing effort in Sapphire to classify scientific data sets.

Automatic Performance Experiment Management and Analysis

Michael McCracken
University of California, San Diego

Summary:

Performance experiments are a common task for researchers at LLNL. The experiment space for these can be characterized by a set of control variables such as number of processors or an environment setting. Usually, there is a small set of metrics in which the experimenter is interested, for example, application runtime or cache utilization. The growing scale of parallel applications and the customizability of runtime systems can result in a very large experiment space that might be interesting to a researcher. The usual way to tackle such an experiment is to change settings, recompile, and run by hand, or to write a custom script that takes care of running the experimental trials but probably cannot assemble the data. This approach is labor intensive and a waste of the researcher's time. Because of the changing nature of the Laboratory's computing systems, a way to automate this common task is required that does not limit the scope of experiments that can be performed.

We have developed a tool, called Sergeant, that is flexible and general enough to manage complicated performance experiments and powerful enough to assist in interpreting the data from these experiments. It allows the description of the experiment space in a few short lines of a configuration file, and relieves the experimenter from the burden of rewriting the program logic that controls the execution of each experiment. The same compact file can describe the data that can be gathered from the experiment, and it takes the work of collating the data away from the experimenter, who is then able to browse only the results that are deemed interesting. Actually running experiments with the Sergeant tool is completely automatic, freeing the researcher to pursue other tasks. Several control space search strategies were investigated to intelligently limit the number of experiments run and to trim the amount of data produced. These strategies included hill-climbing optimization, factorial experimental design and analysis, and correlation analysis.

Use of Sergeant for performance experiments results in a significant decrease in the amount of time it takes to plan and begin such an experiment, a reduction in the amount of redundant programming a researcher must do. It lowers the amount of data that the researcher must search through to make useful conclusions about the experiment. A graphical interface for performance data visualization was developed for Sergeant, which makes the process of analyzing experiment results and producing graphs faster and much more conveniently.

Further development of Sergeant is planned, including investigation of more advanced search strategies and intelligent exploration of the experiment space.

Dynamic Detection of Streams in Memory References

Tushar Mohan
University of Utah

Summary:

With processor speeds doubling every eighteen months and main memory latencies reducing at only seven percent annually, application performance is becoming increasingly memory bound. Computer architects attempt to hide high main memory latencies by adopting techniques such as prefetching and using multi-level cache hierarchies. An application's performance on such architectures depends on the actual memory access patterns exhibited. The ability to recognize streams-sequences of evenly spaced addresses makes it possible to exploit knowledge of future accesses to improve memory system efficiency, as in smart prefetching memory controllers and parallel vector access mechanisms. Earlier work characterized an application's reference patterns using static, compile-time analysis. We present an efficient on-line algorithm to detect streams at run-time by analyzing the memory references issued by the program. By dynamically detecting streams, we avoid the need to create large trace file—issue for all but very short runs of trivial programs. We have implemented our algorithm and have used it for detecting streams in some common codes like matrix multiplication and DAXPY.

Our algorithm detects streams, which we define as a sequence of evenly spaced references. Other stream definitions are possible, but we choose a linear definition because of the widespread occurrence of such streams in real applications. Our algorithm detects streams interleaved with other streams and with irregular references, as well. It does this by computing the stride for a new reference with respect to previous references. On finding a sequence of equal strides greater than a minimum length, the sequence is classified as a stream. We limit the spatial and computational complexity of our algorithm by aging accesses. Older references that are not recognized as part of any stream are discarded as new references are added. We implement this by maintaining a window of active references; thus, streams whose references are further apart than the window size will not be detected. Our algorithm maintains a stream table of detected streams, each of which is succinctly represented as a tuple: starting reference, length, and stride. These tuples are conveniently stored in a chained hash with the expected stream successor as a hash key.

New references are stored in a pool until they can be classified into a stream or discarded through aging. Detecting streams requires determining the differences between a reference and its predecessors in the pool. To reduce the complexity of recomputing these differences as new references are added, we extend the pool to make room for differences to be stored along with the reference. As references age, their differences are retired as well.

The algorithm proceeds by appending new references to the pool. A hash lookup is then performed on the stream table to determine whether the new reference extends an already detected stream. If so, the stream table information is updated and the pool slot is marked to prevent difference calculations for the reference, as well as to ensure its exclusion in subsequent stream detection. If the reference does not extend any stream, its differences with its predecessors in the pool are computed. The differences are then scanned for streams. If a stream is detected, its accesses are removed from the pool and the stream is inserted into the stream table.

Summary (continued):

Tushar Mohan
University of Utah

We have implemented our algorithm and run it on commonly used codes. Though we have used a feature of the compiler on SRC-6 to instrument source codes presently, we are also developing novel techniques for dynamic binary instrumentation to record data references without the need for any compiler support. Currently, the compiler inserts a subroutine call before every load and store, with the address and the type of memory reference passed as parameters to this subroutine. Our algorithm is run as a separate process, and IPC is accomplished through named pipes.

We applied our stream detection algorithm to a few regular codes, like iterative matrix multiplication and DAXPY. We found that our algorithm correctly identifies streams occurring in the references. In matrix multiplication we found the absence of some expected streams and the detection of other streams of longer length. We will be applying our algorithm to a wider variety of codes in the future.

Large-Eddy Simulation and Multigrid Methods

Sandra Naegele
University of Heidelberg

Summary:

Our objective is the use of multigrid solvers for turbulence calculations using Large-Eddy Simulation (LES) modeling. Both multigrid and LES are based on multiscale phenomena and therefore multigrid methods seem well suited for LES calculations. There are some aspects that have to be considered, like the coarse-grid operator, transfer operators, and parallelization, when LES is used in conjunction with multigrid as solver.

The incompressible Navier–Stokes equations are discretized by a finite volume method of second order, and to discretize the system in time a implicit Runge–Kutta Method of second order is used. The unresolved scales of motion, the subgrid scales, are modeled by eddy viscosity type models or mixed models where for both types a dynamic determination of the model parameter is applied. The computational grid is adaptively refined to increase the resolution and decrease the model effort in areas of the domain where complicated structures are located. For the parallelization of the smoother a block-Jacobi approach is applied and the load balance itself is done via recursive coordinate bisection (RCB).

Some simulations were run on the ASCI Blue machine which gave fruitful experience in parallel calculations with regard to Jacobi-effects for the smoother on coarse grids and also with respect to parallelizing the turbulence model.

The main topic will be to use adaptive refinement, and due to that, very unstructured grids for 3D test problems like the flow around a cylinder. The cylinder flow is a very complicated problem where the necessary resolution differs strongly throughout the computational domain. Furthermore, the behavior of the solver as well as different approaches to construct the coarse-grid operator will be investigated.

Modeling Chemical Kinetics in Turbulent Combustion Simulations Using Overture

Diem-Phuong Nguyen
University of Utah

Summary:

The research objective of reaction modeling is to computationally link large, chemical kinetic mechanisms to turbulent combustion computations. The link must incorporate small-scale chemistry into large-scale components of the turbulent flow and systematically reduce the degrees of freedom of the system. The bridging between microscopic details to macroscopic domain is achieved through introduction of a subgrid scale (sgs) reaction model. Thus, my summer research at LLNL involved using the Overture framework to incorporate an sgs reaction model to the OverBlown Navier–Stokes solver to simulate a turbulent open pool fire.

The open pool fire simulation consists of calculating a transient turbulent flow coupled to gas phase kinetics and soot particle dynamics. The OverBlown solver will be used to solve the compressible Navier–Stokes equations along with an additional scalar transport equation for mixture fraction. The CFD is coupled to the reaction kinetics via the mixture fraction and enthalpy. The reaction model then provides the CFD with updated values of the local density, temperature, rates, and state space.

For the open pool fire, the traction free boundary conditions developed by Boersma will be implemented. These conditions require that all domain boundaries except for the outlet adhere to a Neumann condition where the derivatives are set to zero. The outlet adheres to a specified convective boundary equation.

Regarding the fluid dynamics, the Smagorinsky LES turbulence model will be implemented. Molecular diffusion will be neglected and the turbulent diffusion coefficient will be modeled as the turbulent Schmidt number times the momentum diffusivity obtained from the Smagorinsky constant.

In terms of reaction kinetics, two different models are available for coupling to the CFD. They are the equilibrium model and the intrinsic lower dimensional manifold (ILDM) method. Equilibrium consists of gas phase species only while ILDM captures non-equilibrium behavior and incorporates data from 800 reactions and 185 species including soot.

Interface code was written to couple the reaction model module with the OverBlown cfd. The ILDM implementation required an additional interface with a KDTree lookup. Bill Henshaw of LLNL is currently adding LES turbulence modeling, Boersma boundary condition, and scalar transport capability to the OverBlown CFD. Thus, the framework is in place for a series of simulations. This current framework is highly flexible and different reaction models, reactants, and fuels may be used.

I will run the open pool fire simulations and validate the data this fall. I plan to use Overture and OverBlown as the CFD platform to test out different reaction models. I would also like to use the complex geometry capability of Overture to apply turbulent reacting combustion simulations to industrial applications such as reactors. Finally, I would like to investigate the system in a parallel environment.

Space-Time Multigrid

Luke Olson
University of Colorado at Boulder

Summary:

Multigrid methods are considered well-developed and very efficient for elliptic boundary value problems and, depending on the discretization, multigrid methods can also be tailored for non-elliptic problems. In time-dependent problems, the use of semi-discretizations and time-stepping methods, which often result in elliptic problems at each step, are the method of choice. There are potential drawbacks in this framework, however. In large problems, parallelism becomes desirable, yet scalable time-stepping methods are rare and complex since they are inherently sequential in design. Furthermore, adaptivity and local mesh refinement are essential in many applications where a small part of the domain needs higher resolution while other parts of the domain need only coarse grids. In time-dependent problems, the areas in need of refinement not only exist spatially, but also in time. Adaptivity and some amount of parallelism are obtainable, but there is a cost involved. Part of the summer research project was to investigate these costs. Another goal was the study of the types of time-dependent problems that might benefit from a space-time multilevel approach and the discretizations that set up a convenient framework for efficient multigrid algorithms and adaptivity for these applications.

The software package *hypre* (High Performance Preconditioners) provided the necessary tools to solve these sparse linear systems in parallel. The flexibility of *hypre* enabled us to compare the "general" cost of solving a problem in space-time with multigrid versus stepping through time through SMG (semi-coarsening multigrid, a.k.a. Schaffer multigrid), a robust multigrid method for anisotropic arid variable coefficient problems that takes advantage of operator-based interpolation, line-relaxation, and semi-coarsening.

The FOSLS (first-order system least squares) methodology seeks to meet the goals of speed and adaptivity. FOSLS provides a sharp error estimator and in many cases leads to a form for fast multigrid convergence. Also, the growing interest from neutron transport practitioners indicates value in future research in this area.

FOSPACK Programming for Elasticity

Moongyu Park
Purdue University

Summary:

The standard Galerkin procedures for elasticity have many practical difficulties (e.g., the “rocking phenomenon”) as the material tends to become incompressible, i.e., the Poisson ratio tends to 0.5. Many people have attempted to develop alternate approaches that can overcome the difficulties in the incompressible limit. However, these approaches are usually based on mixed formulations that lead to discrete equations that are difficult to solve. Cai, Manteuffel, McCormick, and others developed first-order system least-squares (FOSLS) approaches for which Ruge wrote an implementation (FOSPACK) using algebraic multigrid methods (AMG), which is the main tool of our investigation.

We have tested the accuracy and convergence factor of FOSPACK for elasticity problems on several domains such as square, L-shape, horseshoe shape, rectangular domain with a rectangular hole, circular domain with a circular hole, and a rectangular domain with a circular hole. It has O(h**2) accuracy on square domains and a little lower accuracy on other domains because of the geometry irregularities of the domains.

With these successes demonstrated, we will be working on other model problems and developing more robust and higher performance implementation.

Measuring the Regularity of Array References

Erin Parker
Purdue University

Summary:

The running times of large scientific programs are strongly influenced by the time spent accessing main memory. Many mechanisms, such as prefetching, exploit regular access patterns in order to overlap memory accesses with computation and, thus, reduce memory stall cycles. The benefit of these mechanisms depends on the regularity of an application's memory accesses. Although several access descriptors have been proposed, access regularity is an intuitive concept for which few formal metrics exist.

We consider a program to be regular if it contains array references with identifiable access patterns that are repeated as memory is traversed. For our purposes, we restrict this definition to linear patterns. We present a set of metrics that quantify access regularity. We have implemented a source-to-source compiler mechanism to measure access regularity. Results on our sample code demonstrate that our analysis mechanism is fast and accurate.

We present three approaches for measuring the regularity of a program. Our static approach is a low run-time overhead mechanism that uses statically determined information, augmented at run time only by simple scalar data, such as loop bounds. Our dynamic approach instruments array references so that their regularity can be precisely determined at run time; this approach has significant run-time overhead but is highly accurate. Our overall approach is a hybrid of the static and dynamic approaches. It provides high accuracy with reasonable run-time overhead by using statically determined information where possible.

The static approach examines a program's AST (Abstract Syntax Tree) at compile time to gather knowledge of its loop nests and the array references made within the loop nests. Based on analysis of the array index expressions, we categorize an array reference as regular, irregular, or indeterminate. A regular array reference is one in which all indices are linear expressions of the LCVs (Loop Control Variables). An irregular array reference is one in which at least one index is a nonlinear expression of the LCVs. An array reference is indeterminate if at least one index is an expression that cannot be analyzed at compile time or the array reference is contained in the body of a conditional. For example, the array reference A[B[i]] is indeterminate without knowledge of how array B is initialized, and the array reference A[f(i)] is indeterminate without knowledge of what is returned by the function f given input i. Although more aggressive compile time analysis can categorize some occurrences of these two examples as regular or irregular array references, in general, their regularity cannot be determined until run time.

For any regular array reference, each execution of the innermost loop enclosing it will generate a predictable stream of array accesses. We call such a stream a regular stream. Our linear restriction implies that array references we classify as irregular do not constitute a regular stream. Indeterminate array references may be irregular; our static approach assumes that they are. Therefore, based on analysis of the LCVs of the loop nests containing regular array references, we can compute the number of regular streams, their average length, and the proportion of array accesses that occur in regular streams, among other statistics.

Summary (contiued):

Erin Parker
Purdue University

The dynamic approach examines a program's AST to locate array references contained in loop nests. It does not analyze the indices of array references or LCVs of loop nests. Instead, we instrument the AST with instructions for tracking the actual value of the index to an array reference. A stream of indices forms a regular stream if the stride between all values is the same. We keep the same statistics for regular streams as in the static approach. The dynamic approach accurately categorizes all array references although it makes no attempt to categorize them statically.

The hybrid approach combines the two approaches described above. As in the static approach, we categorize an array reference as regular, irregular, or indeterminate. Then for every regular array reference, we compute the statistics for its regular streams. However, instead of conservatively assuming that every indeterminate array reference is irregular, we perform run-time tracking of array indices to discern actual regularity, as in the dynamic approach.

It is clear that the static approach incurs virtually no run-time overhead, but its accuracy can vary widely and is based on the number of indeterminate array references in a program. The dynamic approach enjoys great accuracy at the cost of a noticeable run-time overhead. The hybrid approach is designed to incur larger run-time overhead only when it is necessary for greater accuracy.

We accomplish automatic analysis and instrumentation of the AST using ROSE. ROSE is a tool for building source-to-source preprocessors. The preprocessor generates an AST from the program source code; the AST is then used for analysis, instrumentation, or optimization. The instrumentation of our static approach merely computes the regularity statistics once the values of any run-time constants are known. In our dynamic approach, our instrumentation actually tracks array index values and detects any regularity. The hybrid approach only uses the more expensive run-time instrumentation for indeterminate references.

We discuss the accuracy and run-time overhead of our three approaches for a simple test program. This example program clearly demonstrates the trade-offs between our approaches.

```
do i = 0, regularity_param

do j = 0, MAX

sum += A[j]

do i = 0, 100 - regularity_param

do j = 0, MAX

sum += A[B[j]]
```

Note that regularity_param is an integer provided by the user at run time whose value is between 0 and 100. B is an array of integers with size at least MAX, which has been initialized in one of two ways. In Case 1, B[i] is a random integer with a value between 0 and MAX-1, and in Case 2, B[i] = i.

Summary (contiued):

Erin Parker
Purdue University

Our example program has three array access streams: the accesses to the A array in both loops and the accesses to the B array that determine the A indices in the second loop. All three of our approaches correctly detect regularity in Case 1 of the sample program. However, for Case 2 of the sample program, our static approach misclassifies array reference A[B[j]] as irregular, while our dynamic and hybrid approaches correctly classify it as regular.

The running time required by the source code instrumented using our hybrid approach is proportional to the number of indeterminate array references that must be tracked at run time, as expected. The run-time overhead of our hybrid approach is significantly less than that of our dynamic approach even when the value of regularity param is 0. Although the indeterminate array reference A[B[j]] must be instrumented for run-time detection of regularity, our hybrid approach saves run-time overhead by statically categorizing the array reference B[j] as regular.

The effort to measure regularity in programs is ongoing, and the preliminary work discussed here has raised several issues. It is undesirable to use our dynamic approach to measure the regularity of large LLNL codes, as it will add overhead to already long-running programs. Likewise, the possible inaccuracy of our static approach on complicated programs makes it unsuitable. Therefore, we are interested to see the accuracy/overhead trade-off of using our hybrid approach on such programs. Furthermore, our analysis can be expanded to include references to array class objects in use at LLNL, which we expect to introduce new challenges.

A Shrink-Wrapping Approach to Large-Scale Visualization

Serban Porumbescu
University of California, Davis

Summary:

Scientists are running simulations that generate data sets several orders of magnitude larger than what current technology can handle. Our goal is to create algorithms that allow interactive manipulation and visualization of these large data sets.

Our technique, called "shrink wrapping," takes a generated isosurface (mesh) and reparameterizes it so that wavelet compression can be applied. This summer we modified our current approach by eliminating the construction and use of the signed distance field due to surface folding problems we could not resolve. Our new approach creates a lower resolution mesh by performing a topology preserving decimation of the initial isosurface. This newly created mesh is then also decimated to create a third mesh. The process continues until we generate a base mesh with a suitably low resolution. At this point we have a continuum of meshes where each mesh differs from its finer and lower resolution version by a very small amount. Beginning with the base mesh, we take each mesh in turn and perform a series of subdividing, snapping, and smoothing operations in relation to the mesh at the next finer level. The idea is to capture important features by gradually introducing them as we move from coarse to fine meshes.

Our future work entails fine-tuning the newly developed technique, parallelizing our algorithm, and applying an adaptive mesh approach that performs the subdivision operation only in areas of high detail.

Simulation Tool For Studying Solutions to the Memory Wall

Pete Poulos
University of Utah

Summary:

Processors are getting much faster than memory systems. These diverging speed growth curves create a memory bottleneck in almost all computer systems, but the problem is particularly acute in symmetric multiprocessor systems (SMPs). The Memory Wall project is developing techniques to improve the performance of the memory hierarchy in SMP systems. This requires exploring design alternatives for smart memory controllers. To this end, we are developing a simulator that will allow us to gather statistics about the performance of the various design alternatives without having to build each candidate in hardware. This will save time and money, as it is easier to simulate a system than it is to build one.

We need to be able to simulate a wide variety of hardware configurations. Also, the simulator needs to run fast and simulate parallel architectures. The latter requirement distinguishes it from most publicly available simulation tools. In order to make the simulator efficient, we have designed it to be easily parallelizable into a number of threads to be run on a parallel computer. The parameters of the machine being modeled and the statistics to be gathered during simulated execution are settable at run-time via a single configuration file.

We are using run-time trace-driven simulation, where each input trace is gathered from a single node of a baseline SMP system. These memory traces are presented to our simulator, which calculates timing information for each access on the memory system being simulated. Using traces allows us to abstract away all details of the processor microarchitecture and instead focus on the memory behavior. This lets us model any processor and increases simulator efficiency, but potentially sacrifices accuracy in measurements. This is an appropriate tradeoff for this particular research project. This type of simulation is not publicly available.

Many of the individual system components have been developed, but we are still implementing the system-wide architecture. The API between the program being traced and the simulator is being developed, as well as the components that manage the division of components across different threads.

Gesture-Tracking for Visual Intuitive Navigation of Large Data Sets

Min C. Shin
University of South Florida

Summary:

We have been developing a new approach to a gesture-tracking system using real-time range on-demand. The system represents a gesture-controlled interface for interactive visual exploration of large data sets. The paper describes a method performing range processing only when necessary and where necessary. Range data is processed only for depth of interest. This is accomplished by a set of filters on the color, motion, and range data. The algorithm also includes a robust skin-color segmentation insensitive to illumination changes.

We have tracked the manipulating hand nearly at four frames per second. By using change of hand area, we have identified the opening (rapid increase of area) and closing (rapid decrease of area) signifying the beginning and ending of the gesture. Using the range images, we have computed the 3D trajectory of the manipulating hand. We fitted the 3D Bezier curve using the 3D trajectory. Then, the curvature and the trajectory direction are used to determine the gesture. If the curvature is high, the gesture is recognized as rotation. If the curvature is low indicating a linear motion, we have classified the gesture as zooming if the direction is mostly in z-axis; otherwise it is classified as translation.

We have tested the algorithm using 398 images of two people consisting of 16 gestures. The algorithm was able to detect the manipulating hand 99% correctly. The beginning and ending of gesture were detected 91% correctly. Overall, the gestures were recognized correctly with an 88% success rate.

We aim to further improve the speed by computing range on selected regions (range on demand). We will also connect the system to a visualization package to demonstrate the application of the gesture tracking.

Parallel Grid Refinement and a Posteriori Error Analysis

Stanimire Z. Tomov
Texas A&M University

Summary:

Parallel grid generation tools play an important role in scientific research. To enable the development of efficient computational technologies, such tools may have to generate finer meshes only in some regions of the computational domain. This could be achieved by applying a posteriori error analysis. The goal of my summer project was to work on this problem area. More precisely, my goal was: (l) to further develop a parallel grid generation tool for three-dimensional problems; (2) to help other researchers in CASC to use it for algorithm testing purposes; (3) to integrate it with *hypre* data structures; (4) to use it with *hypre* preconditioners through the Finite Element Interface (FEI); and (5) work on a posteriori error analysis from theoretical and practical points of view.

A parallel mesh generation tool, named ParaGrid, was further developed. The new development was a continuation of a two-dimensional project begun last summer. ParaGrid is software that takes as input a coarse tetrahedral mesh, which describes well the domain, splits it using the METIS partitioning software, distributes the partitions among the available processors, and generates in parallel a sequence of meshes. It has internal solvers and is able to generate various Finite Element/Finite Volume discretizations. The data structures allow ParaGrid to be easily connected to (or used to provide data to) external parallel finite element/volume solvers based on domain decomposition. It has been successfully used from several researchers in CASC for algorithm testing purposes.

I worked with Charles Tong on data structures for parallel finite element problems. The stress was on the generation of a parallel element topological data structure. I finished the generation of *hypre* matrices by providing the relations "elementnode," "elementface," "facenode," and their transposes. The test data was generated with ParaGrid.

Generation and solution routines for elasticity problems were added to the code. *hypre* preconditioners and solvers can be used. The connection is done through FEI 3.0.

I completed with Dr. Raytcho Lazarov a study on a posteriori error control strategies for finite volume/element approximations of second order elliptic differential equations. These refinement techniques were applied to finite volume discretizations of various boundary value problens for steady-state convection-diffusion-reaction equations in two and three dimensions. The results were summarized in an article on "A posteriori error estimates for finite volume element approximations of convection-diffusion-reaction equations," which has been submitted to *Comput. Geosciences*.

An accompanying visualization a tool named GLVis was also developed. GLVis started from a 2D visualizer developed in a team project at Texas A&M University. Its features include solution visualization in moving cutting planes, input from files and AFJNET sockets, visualization of vector field, and displacements.

My work on error control, adaptive grid refinement, and a posteriori error analysis continues. In addition, mesh de-refinement software is under development for the case of adaptive grid refinement for time-dependent problems.

Input/Output Scalability of Different Architectures

Preethy Vaidyanathan
University of California, Santa Cruz

Summary:

We seek to understand how to develop file systems for low-cost cluster computers for biological applications. Many studies have been carried out with file systems for classical parallel applications (e.g., PVFS). Like other scientific computations, many computational biology applications are highly data parallel, but they have the distinct characteristic of being I/O intense.

We instrumented an application vital to the Human Genome Project, which consumes more than half of the execution time at the UCSC computational biology cluster using the Pablo Instrumentation library from the University of Illinois. We characterized the I/O behavior of this application on Linux clusters with different file systems (the UCSC cluster and the Vivid cluster) and an IBM SP/2 (ASCI Pacific Blue). Our study shows that disk-processor locality is a very important factor affecting the performance of this application on the cluster.

We have presented the design of a user-level library for a new model of location-transparent storage to automatically redirect read accesses to the most appropriate location for obtaining the best performance.

The Discretization of the Scattering Kernel with Angular Finite Elements

Nicolas Vallete
Texas A&M University

Summary:

The angular variables in the transport equation are usually discretized either by a spherical harmonic expansion or by a quadrature rule (so-called "Discrete Ordinates" methods). Discrete Ordinate methods often yield unphysical results, which are called ray effects. The goal of my summer project was to reduce the ray effects by discretizing the sphere with finite elements. The discretisation of the transport equation by piecewise constant finite elements leads to an equation that can be solved with existing algorithms.

The goals for the summer were to code the finite element representation of the scattering kernel, determine the underlying symmetry of the scattering kernel, code the symmetries into the function, which calculates the scattering kernel, and to implement the previous program into a parallel transport code in order to compare this method with others.

The initial program generating the integrals of the spherical harmonics was written in Matlab. I completed this program with some other subroutines in order to generate cross sections tables. I also wrote another program in Matlab, which reduces the amount of storage of the scattering cross section by a factor of 50. Future plans for this method include adaptive refinement on the sphere.

Diffusion Synthetic Acceleration for Three-Dimensional Transport Equations

S. Van Criekingen
Northwestern University

Summary:

The linear Boltzmann transport equation (BTE) is an integro-differential equation arising in deterministic models of neutral and charged particle transport. Iterative methods are routinely used to solve large equation systems resulting from discretization of the BTE. The diffusion synthetic acceleration (DSA) has been proved to be very efficient in preconditioning the one-dimensional BTE. Our aim was to extend existing one-dimensional convergence proofs to three dimensions.

The discretizations consist respectively of a standard discrete ordinates collocation of the angular variable, and a Petrov–Galerkin finite element approximation in space. DSA requires a "consistent" discretization of a limiting diffusion approximation to the BTE. While for 1D slab geometry the consistently differenced diffusion problem is nonsingular, it has been shown that the consistently differenced 3D diffusion approximation is actually singular, although the DSA preconditioner itself remains nonsingular. We extended the 1D theoretical results in the asymptotic diffusion limit to 3D and submitted our work to *SIAM Journal of Numerical Analysis*.

Additional work is needed to address the DSA method in the context of the corner balance spatial discretization method, for which some 1D results have been obtained already.

Adaptive Mesh and Algorithmic Refinement Simulations for Multiscale Hydrodynamics

Sanith Wijesinghe
Massachusetts Institute of Technology

Summary:

LLNL has a multifaceted research effort focusing on numerical, algorithmic, and software issues related to the use of structured adaptive mesh and algorithmic refinement (AMAR) technology. A key benefit of AMAR is the capability to investigate multiscale and multiphysics problems within a single computational environment. This feature is key to the analysis of many computational physics applications (e.g., Richtmyer–Meshkov instability) important to LLNL and the ASCI program.

The AMAR scheme used in the current research consists of a second-order Godunov Euler solver and a Direct Simulation Monte Carlo (DSMC) particle solver.

The objectives of the current research are as follows:

1) Assess AMAR simulation through extensive testing of model problems.

2) Investigate methods to track fluid discontinuities and material interfaces and how they impact the results obtained from AMAR.

3) Scale AMAR to large systems of interest.

Test simulations for stationary shock waves have been conducted. The results from these simulations are currently being compared with theoretical results. Concentration-based adaptive gridding criteria have been incorporated to track material interfaces. AMAR simulations of a step concentration discontinuity have validated these criteria. A large-scale parallel simulation of the Richtmyer–Meshkov instability has been conducted on the IBM ASCI Blue.

Future plans include improvements to the load balancing routines to allow more efficient parallelization of the AMAR scheme. Longer Richtmyer–Meshkov simulations will also be conducted to investigate the time evolution of the material interface.

Multiscale Simulation Combining Direct Simulation Monte Carlo and Navier–Stokes Solvers

Yihao Zheng
University of California, Davis

Summary:

Numerical modeling of fluids is particularly challenging when the problem of interest spans length scales differing by orders of magnitude. Even mesh refinement is not sufficient when the smallest length scale approaches the microscopic regime since standard hydrodynamics is not accurate. In such cases a different simulation algorithm, which contains the appropriate microscopic physics, is required but only at the finest level of refinement; this methodology is known as Adaptive Mesh and Algorithm Refinement (AMAR).

These types of hybrids, combining microscopic and macroscopic algorithms, are still under development and many critical challenges exist, yet their utility is well recognized in the field of computational fluid mechanics. Our specific project involves the implementation of such a hybrid code combing Direct Simulation Monte Carlo (DSMC) for particle simulation and a Godunov-type Navier–Stokes solver for continuum modeling, using the Structured Adaptive Mesh Refinement Application Infrastructure (SAMRAI).

The hybrid code combines two simulation elements, one of which (DSMC) was already in place at the start of the summer. The other half is a Navier–Stokes solver, which is being developed as a generalization and extension of the Euler solver currently used by the AMAR hybrid code. The work this summer consisted of reading the supporting literature and learning the existing code. A stand-alone version of the three-dimensional, second-order Godunov solver for the Euler equations was written and is now being tested.

Future plan call for generalizing the solver to the Navier–Stokes equations and integrating the new solver into the existing AMAR code.

Institute for Scientific Computing Research

Workshop and Conference Reports

Synopsis of Workshop Events:

Bay Area Scientific Computing Day

The Bay Area Scientific Computing Day is an annual gathering to encourage the interaction and collaboration of researchers in the field of scientific computing from the Bay Area. It also provides researchers who are new to the Bay Area an opportunity to present their work to the local community. BASCD 2001 was held on Saturday, February 10, and was hosted by Lawrence Livermore National Laboratory and organized by CASC and the ISCR. More than 120 researchers attended the event, from laboratories, universities, and industry. Thirteen technical presentations were given by graduate students, postdoctoral researchers, and other scientists nominated or invited from throughout the Bay Area.

Program details are available at the BASCD 2001 website located at: http://www.llnl.gov/CASC/workshops/bascd/

Synopsis of Workshop Events:

Workshop on Object-Oriented and Component Technology for Scientific Computing

The first Workshop on Object-Oriented and Component Technology for Scientific Computing was held July 23-25 in Livermore, California. Forty-six active researchers and developers of scientific software discussed the state of the art in the field of high-performance software technology. Participants came from a variety of backgrounds, including thirty-one from the DOE, three from NASA, eight from universities, and four from industry.

This workshop was divided into five sessions focusing on software issues that arise in scientific computing: scripting, component technology, distributed computing, problem solving environments (PSEs), and applications and libraries. Scripting provides a flexible and interactive mechanism to couple and steer large science simulations. Component technology provides a means to manage the complexity of modern scientific simulation software and enables new simulation capabilities that were previously unavailable due to limitations in interoperability. Distributed computing frameworks provide the illusion of homogeneity in a heterogeneous environment. Problem-solving environments provide rich functionality for development and deployment of domain specific software. Applications and libraries are the final products (and from the point of view of component technology developers) for the customers of these technologies. Each of these five sessions consisted of four or five speakers. A one-hour moderated panel discussion covering the general topic of the session followed.

This format worked exceedingly well because questions for the panel were submitted in advance when participants registered. These “killer questions” were advertised on the workshop’s website, and the speakers were reminded that if they did not address these questions in their talks, they would almost certainly have to do so in the panel discussions that followed.

Among the attendees, there quickly grew up a common theme about the difficulty of software development and integration. Very few speakers discussed their software in isolation. Most talks highlighted either how their software could be integrated into other packages, or how they used other packages to augment their own software.

It became clear that software reuse is a critical necessity for scientific computing to continue to advance. To meet the push for more physics, higher fidelity, and more functionality software has to become more reusable… and usable in broader contexts.

Workshop participants identified three major barriers to effective software reuse: interoperability, semantics, and portability.

- Software interoperability is hindered by language incompatibilities, the lack of standard scientific interfaces, failure to adhere to existing standards, software build and installation incompatibilities, and too strong assumptions much about the local computational environment.

Synopsis of Workshop Events (continued):

- Participants agreed that not enough practical research was being done in the area of scientific software semantics. Most current compilers and source-to-source translaters specify only the calling syntax of a scientific library. Semantics descriptions would extend that syntactic description to specify the behavior and valid usage of a software library.

- Finally, participants identified portability as a difficult and time-consuming task. Portability involves building software and the associated compilation system that is robust enough to work on multiple computer platforms. Portability is especially a concern in high-performance computing, since parallel platforms have a much shorter lifespan than the software, and the operating systems for most parallel platforms are specifically modified for the parallel programming environment.

The timing of this workshop was auspicious, as it came on the heels of the SciDAC awards and provided a showcase for CASC's role in developing Component Technology for scientific computing.

Program details are located at the conference website at http://www.llnl.gov/CASC/workshops/components_2001/

Synopsis of Workshop Events:

14th Copper Mountain Conference on Multigrid Methods

The 14th Copper Mountain Conference on Multigrid Methods was held on April 1-6, 2001. Over one hundred applied mathematicians and computer scientists from eleven countries attended the meeting and over thirty graduate students and post-docs were supported. During the five-day meeting, 68 talks on current research topics were presented, as well as three tutorials preceding the reception on Sunday.

Presentations with similar content were organized into sessions, including:

- Transport
- Inverse problems
- Algebraic methods
- Domain decomposition
- Nonlinear methods
- Fluid dynamics
- Ocean dynamics
- Multilevel theory
- High-performance computational strategies
- First-order systems of least squares

In addition, a lively Circus (of extemporaneous presentations, in the tradition of the Finite Element Circus) was convened Thursday night. All of the sessions, including the Circus and the tutorials, were very well attended. One of the hallmarks of the Copper conference series is the relaxed atmosphere and open, active discussions that it fosters. This collaborative environment was again very much in evidence at this year's meeting.

A student paper competition stimulated student participation in the Conference. This year, there were 25 students in attendance and about 15 returning scientists who first attended the Copper conference series as students. For the competition, students were asked to submit a singly authored paper containing original research. A panel of judges made up of members of the Program Committee selected four winners. These winners were: Avraham Kenigsberg, Malik Silva, Chisup Kim, and Markus Korwarschik.

Selected papers from the conference will appear following a standard review process in E.T.N.A.

Program details are located at the conference website at http://amath.colorado.edu/faculty/copper/2001/

Synopsis of Workshop Events:

The Third Workshop on Mining Scientific Datasets

The Third Workshop on Mining Scientific Datasets was held April 7, 2001, in conjunction with the First SIAM International Conference on Data Mining. The goal of the workshop was to bring together researchers from the data mining community and various science and engineering communities in order to better understand how data mining can be used for the exploration of scientific datasets. In particular, we hoped to identify the common threads across the diverse technical areas that could be effectively harnessed to solve the problems of scientific data analysis.

Unlike the previous two workshops in the series, where many of the talks were given by invited speakers, we solicited short papers on relevant subjects for the third workshop. These were reviewed by the workshop organizing and program committees. Of the four workshops that were held on the final day of the conference, the scientific data mining one had the largest number of registered attendees. Though the workshop was held on a Saturday, most of the 75 attendees remained throughout.

Each of the morning and afternoon sessions started with a keynote talk. The first keynote talk was by Dennis DeCoste from JPL, who described the work being done at NASA on "Mining Large Datasets using Support Vector Machines." The second keynote talk of the workshop was given by Sara Graves from the University of Alabama at Huntsville on "Creating an Environment for Scientific Data Mining." She described the ADAM toolkit that has been developed at UAH for remote sensing data.

These keynote talks complemented the contributed talks very well. There were 14 contributed talks, which covered a diverse set of subjects including mining astronomical data, multiresolution representation of structural mechanics data, use of boosting techniques to classify remotely sensed data, archiving and processing of remotely sensed data, and the mining of corn-field data to detect insurance fraud. Several of the talks were presented by students on their ongoing research work. During the breaks, it was interesting to see the astronomers exchanging ideas with the remote sensing data analysts about the use of a common set of tools for both "look up" and "look down" data.

The workshop was supported by the Army High Performance Computing Research Center and the Center for Applied Scientific Computing at the Lawrence Livermore National Laboratory. The latter provided funds for travel for eleven students and the keynote speakers. The co-organizers were Michael Burl (NASA Jet Propulsion Laboratory), Chandrika Kamath (Lawrence Livermore Laboratory), Vipin Kumar (University of Minnesota and AHPCRC), and Raju Namburu (Army Research Laboratory).

A web page with details on the workshop is at http://www.ahpcrc.umn.edu/conferences/

Synopsis of Workshop Events:

The 2001 International Conference on Preconditioning Techniques for Large Sparse Matrix Problems in Industrial Applications

The 2001 International Conference on Preconditioning Techniques for Large Sparse Matrix Problems in Industrial Applications (Preconditioning 2001) was the second conference of its kind to focus on preconditioning techniques for solving various sparse matrix problems. The first conference, which was viewed by the community as a big success, was held at the University of Minnesota, Minneapolis, in June 1999.

The Preconditioning 2001 Conference was held at the Granlibakken Conference Center in Tahoe City, California, on April 29 - May 1, 2001. About seventy participants attended the conference; many were leading experts in the area of preconditioning techniques. The conference featured nine invited plenary presentations, twenty-three contributed talks, and fourteen poster presentations. Overall, the presentations were of high quality, and there was much interaction among the conference attendees.

The attendees came from academia, research laboratories, and industries. The U.S. Department of Energy (DOE) was well represented at the conference; there were participants from six of the major DOE research laboratories (Argonne, Lawrence Berkeley, Lawrence Livermore, Los Alamos, Oak Ridge National Laboratory, and Sandia). Some participants were from research laboratories overseas, including Japan, France, and the United Kingdom. Participants from industry included Bell Labs, Boeing, Chevron, Fujitsu, Schlumberger, and Xerox PARC. Academic participants came from universities in the United States, Belgium, France, Germany, the Netherlands, Spain, Switzerland, Sweden, and the United Kingdom.

The conference received support from DOE Headquarters, Lawrence Berkeley, Lawrence Livermore, Sandia, and the University of Waterloo. The conference was also sponsored by the SIAM Activity Group on Linear Algebra.

It is well known that the kernel of many large numerical simulations is the solution of large sparse matrix problems, including in particular the solution of systems of linear equations and the solution of eigenvalue problems. It is not uncommon for the solution of matrix problems to dominate the cost of the entire simulation. Efficient solution of matrix problems therefore has become extremely crucial and will make a big impact on large-scale scientific and engineering simulations.

Examples of such simulations include petroleum refinery processes, environmental modeling (e.g., global warming, groundwater contaminations), structural dynamics, finance, design of advanced semiconductor devices and circuits, accelerator design, fusion energy systems, and aerodynamics.

The traditional means of solving sparse matrix problems is direct factorization. For matrices of small to medium sizes, direct approaches are often adequate. However, as demand for high resolution in the simulations increases, the size of the matrix problems also increases. For extremely large sparse matrices, direct approaches are no longer viable as the amount of memory required can vastly exceed the memory that is otherwise required to represent the application, particularly for three-dimensional simulations. Even if

Synopsis of Workshop Events (continued):

enough memory were not the bottleneck, the time required to solve such large matrix problems would become prohibitive.

An alternative is to use iterative methods to solve these extremely large sparse matrix problems, and research effort has been poured into powerful and efficient iterative methods for decades. However, known low storage iterative methods alone are not sufficient since they lack robustness or require too many iterations; preconditioning techniques are necessary to improve convergence.

Robust preconditioning techniques is a relatively young field when compared to basic direct methods and iterative methods. Many of the techniques developed during the early years did not have much impact initially because of their simplicity and also because of the relatively small size of the matrices to be solved. However, more and more computational experience indicates that a good preconditioner holds the key to an effective iterative method. The big impact of these simple techniques on the performance of an iterative method has attracted increased attentions in recent years. Parallel computers also generate many new research topics in the study of preconditioning. Many new promising techniques have been reported.

However, the theoretical basis for high performance preconditioners is still not well understood; many existing techniques still suffer from lack of robustness; promising ideas need to be tested in real applications. This was the motivation for holding a conference specifically dedicated to the issues in preconditioning and industrial applications. Researchers and scientists/engineers in this field from academia, industries, and research laboratories were brought together so that they could present the progress, discuss difficult issues, exchange their findings, and explore possible new directions.

The topics covered at the Preconditioning 2001 Conference included:

- Incomplete factorization preconditioners
- Domain decomposition preconditioners
- Approximate inverse preconditioners
- Support graph preconditioners
- Multi-level preconditioners
- Preconditioning in eigenanalysis
- Preconditioning in optimization
- Preconditioning in finite element applications
- Preconditioning in image processing
- Applications in fluid dynamics and magnetohydrodynamics
- Applications in multiphase subsurface flow

The conference program included invited plenary lectures, contributed talks, and poster presentations. Selected papers presented from Preconditioning 2001 will appear following a standard review process in a special issue of Numerical Linear Algebra with Applications.

Program details are located at the conference website at http://www.nersc.gov/conferences/pc2001/

University of California
Lawrence Livermore National Laboratory
Technical Information Department
Livermore, CA 94551